A SPATIAL ANALYSIS OF CORN AND SOYBEAN YIELDS AND WEATHER RELATIONS

BY

DI YANG

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricultural and Applied Economics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Adviser:

Professor Scott Irwin

ABSTRACT

This thesis studied the relationship between corn and soybean yields and weather conditions in U.S. Corn Belt area. Three statistical models were tested, including two fixed effect models and one geographically weighted regression. Due to different underlying assumptions and specifications of each model, results indicated different implications. Major findings included severe weather conditions during plant's reproductive stage had much greater impact than vegetative stage; the favorable effect of moderate heat could not compensate the yield loss caused by extreme heat at the same magnitude; and heterogeneous crop-weather relations across crop reporting district prevented further aggregation of cropping region. As another objective of this thesis, predictive powers of each model were also studied.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

1.1 Background

The study of determinants of crop yields has existed for a long time, and is still a topic of great interest nowadays. Not only because sufficient crop production is needed to guarantee enough food supply for human beings around the world, but also agricultural production has become an important element as a feedstock for energy.

Previous studies formed two main categories: agronomy simulation models and empirical statistical models. Simulation models quantify various stages of plant physiology and involve highly complex mathematical models. For instance, the recent Hybrid-Maize model (Yang et al., 2004) for corn yield adopted explicit functions for photosynthesis and respiration, and created maize-specific formula for phenological development and organ growth. To forecast potential crop yield precisely, simulation models often required many specific inputs such as crop type, planting density, and so forth. Therefore, predictions of crop yield only apply to targeted areas, and macro-level forecasts, for example, on the state level, are infeasible.

Empirical studies, on the other hand, focused on general patterns in crop yield determinants. Thompson (1962, 1963, 1969, 1970, 1985, 1986, and 1988), as a pioneering researcher, developed a multiple regression framework to explain weather, technology, and crop production relationships. The main factors studied were monthly average temperature, monthly average precipitation, and trend variables as representations for technology change. Recent research (Roberts and Schlenker, 2010, 2011; Schlenker and Roberts, 2006, 2008, 2009) also attempted non-linear models to explore variations in crop yields. The heat effect was measured by growing degree days, which was argued to be better than average temperature since the extreme heat would not be diluted. All in all, empirical studies emphasized statistical evidence from the data, but may overlook necessary information from plant physiology.

1.2 Objectives

This thesis serves three main purposes. As an empirical study, one objective is to develop crop-weather models that will combine advantages from both agronomic studies and earlier statistical models and avoid their shortcomings. Three models will be developed in this study, each one with specific assumptions and specifications. The crops studied in this thesis are corn and soybeans in the U.S. Corn Belt area.

Yield forecasting has always been an important topic for farmers, hedgers, consumers, and so forth. *World Agricultural Supply and Demand Estimates* (WASDE) from U.S. Department of Agriculture (USDA) is published monthly with predictions of U.S. crop productions since 1973. Previous studies (for example, Irwin et al., 2008, 2009a, 2009b) have also attempted to make crop yield forecasts. With new crop-weather models, another objective of this thesis is to produce more precise corn and soybean yield forecasts.

The last objective is to compare forecasting performance among models. Since the three models have different underlying assumptions and specifications, one interesting topic is to find out the model with the best predictive power. A comprehensive competition will be performed to identify models with the most accurate predictions.

1.3 Data and Method

Three major sources for crop yields and weather data are the National Agricultural Statistics Service of USDA, National Climatic Data Center (NCDC) and data set prepared by

Michael Roberts from North Carolina State University and Wolfram Schlenker from Columbia University. Corn and soybean yields and planting progress data were collected from 1960 to 2010 on a crop reporting district (CRD) level. Since some crops in Nebraska were heavily irrigated, data from non-irrigated yields were used for those CRDs. Monthly precipitation and monthly temperature data collected from NCDC were on a climate division level, which do not necessarily coincide with the geographical definition of USDA crop reporting district boundaries. Thus, a few modifications were made to match data from different sources. Finally, growing degree days as a measurement of heat effect on crop growth were also adopted. However, these data were only available up to 2005, so the analysis was limited.

To explain effects of weather on crop yields, three statistical frameworks were employed in this thesis. Thompson and Roberts and Schlenker models were modified by adding fixed effects. Modified Thompson model evolved gradually through a few researchers, and one major improvement from this study is the introduction of agronomic concepts into regression analysis. Specifically, the temperature effects during vegetative and reproductive stages were treated separately in the model. On the contrary, in the modified Roberts and Schlenker model, one basic assumption was that the temperature effect on yield is cumulative, and therefore additive and substitutable over time. In addition, precipitation was measured with total rainfall during the growing season instead of monthly precipitation. The third model was a panel geographically weighted regression (GWR) model. While the first two fixed effect models only allow the intercept to vary for different CRDs, spatial heterogeneity is captured in the panel GWR model by allowing every CRD to have their own coefficients. The functional forms were in conformity with modified Thompson models so that the analysis will include data up to 2010.

To test and compare predictive power of each model, a forecasting competition was performed using two stragegies, i.e. the recursive method and resampling. In a conventional recursive forecast, new observations will be added one at a time to make new predictions each year from 1988 to 2010. For the resampling method, a sub-sample will be randomly chosen on whole years, and predictions will be made for the rest of the data using results from the sub-sample. In the interest of predicting U.S. total crop yield, two kinds of predictions were calculated. The first is U.S. corn and soybean yield derived from harvested acreage-weighted average forecasts, and the second one is simply the average yield from all CRDs. Composite forecasts of different models were also calculated whenever applicable.

1.4 Overview

This thesis starts with a literature review in Chapter 2. Two streams of study will be reviewed: agronomy theory and empirical studies. While agronomists emphasize the importance of plant physiology and phenology, agricultural economists value statistical evidence from yield and weather data. Chapter 3 provides a summary and descriptive analysis of data sets being used in this study. Most data were available up to 2010 except for the growing degree days data, which were not available after 2005. Chapter 4 illustrates three statistical models in detail, i.e. modified Thompson model, modified Roberts and Schlenker model, and panel geographically weighted regression model. Estimation results of these three models will be presented in Chapter 5. To evaluate forecasting performance of each model, Chapter 6 assesses a comprehensive comparison of predictive power among three models. Finally, a summary of findings and concluding remarks will be given in Chapter 7.

2. REVIEW OF THE LITERATURE

2.1 Introduction

The study of crop yield has always been an important topic as it directly relates to human food supply. Studies with different approaches from various fields have been exercised intensively. On the one hand, agronomists conduct their study based on plant physiology and phenology, which contributes to the understanding of crop yield from a micro level. On the other hand, agricultural economists develop simplified statistical models using generalized weather variables, which serve as macro level research about crop yield. This chapter provides a literature review of some important studies from both perspectives.

2.2 Agronomy Theory

Agronomists use crop simulation models to capture the plant growth processes given certain environment and crop management style. These models generally quantify different phases of plant development and simplify the processes to obtain highly complex mathematical representations. Such studies often require many field experiments and precise understanding of plant physiology. However, since they mainly focus on the effect of certain conditions on a specific development phase of a plant, they may not provide an accurate representation of many conditions on aggregate plant growth and yield. Therefore, sometimes their forecasting ability may not be as good as empirical statistical models.

Among corn simulations models, the Hybrid-Maize model developed by a group of crop scientists at University of Nebraska combines the advantages of generic crop models and maize-specific models (Yang et al., 2004). Generic crop models first describe the plant development and growth without regard to crop species, and then are modified to capture the phenological and physiological attributes of specific crop type. Examples of generic crop models include STICS (Brisson et al., 2003) and INTERCOM (van Ittersum et al., 2003). In contrast, maize-specific simulation models distinguish key development stages of corn and apply different theoretical frameworks rather than generic crop models. For instance, in the maize-specific model MSB (Muchow et al., 1990), temperature is the primary driver of organ growth, and dry matter production is computed directly from absorbed solar radiation; whereas in generic models such as WOFOST (van Diepen et al., 1989), plant organ growth is driven by the availability of assimilates from simulation of canopy photosynthesis, and dry matter production is determined by both growth and maintenance respiration. Hybrid-Maize adopts explicit functions for photosynthesis and respiration used in generic crop models INTERCOM and WOFOST, and embraces revised CERES-Maize formulations for phenological development and organ growth.

In terms of soybean simulation models, SoySim is among one of the most up-to-date models (Setiyono et al., 2010). SoySim incorporates existing approaches for simulation of photosynthesis, biomass accumulation and partitioning, and adds new methodology to simulate flowering, leaf area index, integration of canopy photosynthesis and yield simulation. Main contributions of SoySim include that it only requires easily accessible cultivar-specific parameters and provides validation in high-yield environments which had been missed in existing models such as Sinclair-Soybean (Sinclair, 1986) and CROPGRO-Soybean (Boote et al., 1998).

With simulated crop yield derived from the Hybrid-Maize model, the relationship between temperature and corn yield can be revealed. Generally speaking, corn plant growth can be divided into two main stages: vegetative and reproductive stages. During vegetative stages, the plant uses most resources to grow, and there are emergence, first leaf, second leaf, and so forth to nth leaf, and tasseling stages. Entering reproductive stages, the plant begins to fill the grain that it grows earlier, and there are silking, blister, milk, dough, dent, and physiological maturity stages. While low mean temperatures reduce both photosynthetic rates and kernel-growth rates, it also lengthen the post-silking growth period and grain-filling duration. However, high mean temperatures accelerate crop development and shorten growth duration, and therefore reduce cumulative solar radiation and grain yields (Cassman et al., 2010). In a follow up paper by Grassini et al. (2009), they found that either too high (≈>25˚C) or too low (≈<20˚C) mean daily temperature during post-silking, i.e. graining filling period, will reduce the yield potential.

Similarly, the influence of temperature on soybean growth is identified. Like corn, the growth of soybeans can be divided into vegetative and reproductive stages. The vegetative stages of soybean are very similar to those of corn: emergence, unrolled unifoliolate leaves, first trifoliolate, second trifoliolate, and so forth to nth trifoliolate stages. Nonetheless, the reproductive stages of soybean begin with flowering and include beginning bloom, full bloom, beginning pod, full pod, beginning seed, full seed, beginning maturity, full maturity stages. High temperature during post-flowering period will hasten crop development and shorten seed filling duration, which results in underdeveloped, small seeds. Although seed filling rate can be stimulated by warm condition, the total effect of high mean temperature is still dominated by shortened seed filling duration (Thuzar et al., 2010).

2.3 Empirical Study

The empirical study of crop weather models can be traced back to Smith (1914). It is the first published paper that uses a statistical model to explain the effect of weather on the yield of corn in the central United States (Tannura, 2007). The study concluded that rainfall was the controlling weather factor and July rainfall had the greatest effect on corn yield compared to any other month.

Among all the pioneering studies, Thompson conducted a series of papers establishing a multiple regression framework to explain the relationship between weather, technology, and crop production (1962, 1963, 1969, 1970, 1985, 1986, and 1988). In his papers published in 1985 and 1986, he explored climatic change and corn and soybean production in the Corn Belt area such as Illinois, Indiana, Iowa, Missouri, and Ohio from 1930 to 1984. The statistical model included two major categories of independent variables, i.e. time trend and weather factors. For both corn and soybean production, Thompson noticed significant change in average rates of yield in 1960 due to large fertilizer applications and in 1973 with increased weather variability, resulting in three time trends, each representing one phase. Weather variables used were preseason precipitation (total rainfall from September through June), June temperature, July rainfall, July temperature, August rainfall, August temperature, and their corresponding squared terms.

The results of the corn model revealed that three favorable factors for better corn yields were a cooling trend, greater July and August precipitation, and less weather variability. In addition, three trend variables showed that the use of fertilizers after 1960 had led to a higher rate of yield increase than the previous period, and after 1972 greater weather variability slowed the rate down again. Similar conclusions were found in soybean model as well. Although soybeans did not benefit directly from the fertilizers, they will receive residuals when planted in rotation with corn. However, the trend effect was not as significant as in corn model.

In recent years, Roberts and Schlenker composed several papers exploring the non-linear relationship between weather and crop yields (Roberts and Schlenker, 2010, 2011; Schlenker and

Roberts, 2006, 2008, 2009). Their research was based on a large fine-scale weather data set with county level crop yields in the U.S from 1950 to 2005. To construct the weather data set, they employed the monthly Parameter-elevation Regressions on Independent Slopes Model (PRISM) data and paired them with daily weather station data. Then, they approximated the distribution of daily temperatures and derived the area-weighted average over all 2.5x2.5 mile PRISM grids in a county.

The non-linear regression model was established on an important assumption that the impact of temperature on yields is cumulative over time and yield is proportional to total exposure. This time separability characteristic implies the effect of temperature is additive and substitutable over time. They argued time separability was partially rooted in agronomy, and showed a statistically significant relationship between the cumulative distribution of temperature and yields to validate this assumption implicitly. Nonetheless, they pointed out that in crop simulation models, temperature has different effects on plants during different phases. Total precipitation and technology change represented by a time trend variable also entered the regression model. In addition, county heterogeneity was captured by a time-invariant fixed effect intercept.

The regression results were estimated using three different functional forms of temperature effects: step function, Chebyshev polynomials, and a piecewise linear specification. All specifications showed similar outcomes that temperature contributes positively to yields with a modest rate up to a critical temperature, but reduces the yield considerably beyond that point. They found the critical temperature for corn and soybeans was 29˚C and 30˚C, respectively. The impact of precipitation was statistically significant and quadratic in form. The yield-maximizing rainfall level was found at 25.0 inches and 27.2 inches for corn and soybean, respectively.

Although time trend and fixed effect controls were significant, detailed results were not reported since they were not the focus of the study. Compared with previous statistical models using root mean squared error (RMSE) of out-of-sample predictions, all three specifications outperformed alternative ones. Taking the model without weather variables as a baseline, they concluded new models reduced RMSE by as much as 360% more than alternative specifications and the difference between them was statistically significant.

Based on results from non-linear regression models, Roberts and Schlenker (2010) developed a parsimonious model capturing the effect of extreme heat (degree days above 29˚C and 30˚C for corn and soybeans, respectively), moderate temperatures (degree days 10 to 29˚C and 10 to 30˚C for corn and soybeans, respectively), and precipitation. With a few variations in specification, they concluded that one single best predictor of corn and soybean yields in the United States is extreme heat, and heat and drought tolerance of crops has not improved over the period of 1950 to 2005. The results were similar between time series models and cross sectional models, meaning no significant improvement of heat tolerance has been found either in year-to-year adaptions to weather shocks or in different places that have different growing conditions.

Yu and Babcock (2010) analyzed drought tolerance of U.S. corn and soybeans using data from 1980 to 2008. They created a drought index for each county and regressed crop yields on the drought index and a time trend variable. To measure hotness and dryness of weather, the drought index was a composite representation of temperature and precipitation. It was defined as the product of the number of deviations away from the mean temperature and rainfall, so the higher the temperature or the lower the rainfall, the larger will be the drought index. They stated three advantages of this index. First, compared with state or country level data, the drought index was constructed on county level. Second, this one single index provided an easy means to evaluate two major yield loss causes for corn and soybean: excess heat and deficit of water. Finally, the drought index was highly correlated with yield deviations and identified major droughts years.

The fixed-effect regression model has four specifications. The first specification includes a fixed-effect parameter for each county, a CRD-specific linear trend, the drought index and its squared form and their interaction with time trend terms. The second specification has an exponential trend and only differs from the first one in that the dependent variable was not simply crop yield but the natural logarithm of crop yield, all the right-hand-side variables remaining the same. Then, this log-linear model measures the percentage change in crop yield. The last two specifications were variations from the first two models and were CRD-specific models. They have CRD regional dummy interacting with each independent variable, so all the coefficients are CRD specific.

The estimation results from the linear and the log-linear model showed that for both corn and soybean the drought index was negative and significant and the squared term was positive and significant, indicating the drought had adverse impact on crop yields, and the marginal yield loss and marginal percentage loss decreased with drought severity. The interaction term of drought index and the trend was positive in all models and only insignificant in the linear model for soybean, meaning corn and soybean yields were less vulnerable to minor droughts in percentage terms over time, but susceptibility of soybean to droughts had not changed much compared to the past. The interaction term of squared drought index and the trend was negative and significant in corn models, implying corn yield losses and percentage losses declined under severe droughts, but the magnitude of soybean yield losses and percentage losses were similar under severe and minor droughts. Based on the linear and the log-linear model results, they tested four null hypotheses about yield risk induced by drought. For corn, the test results indicated that yield losses have declined over time no matter whether measured in absolute quantity terms or in percentage terms of mean yields. For soybean, they found the percentage yield losses have decreased over time, but the null hypothesis of absolute yield losses remaining constant cannot be rejected.

The marginal effects of drought were evaluated at the average drought severity level of 1.0. For example, the marginal effect on corn yields declined from 26 bushels per acre or 32% in 1980 to about 18 bushels per acre or 12.4% in 2008, with standard errors of one bushel per acre or 1%. The marginal effect on soybean yields measured in percentage terms decreased from 15.1% in 1980 to 10.5% in 2008, but the absolute losses stayed at about 4 bushels per acre in all years. These findings are consistent with the results from hypothesis testing.

The CRD-specific models generally showed similar results with the aggregate models. However, since each CRD may differ in pre-drought soil moisture conditions, which is unavailable information, the difference in estimates cannot be totally explained. Besides, CRD-specific model only utilized information within a CRD, so this may lead to overfitting problem and large standard errors in estimates. Detailed results of CRD-specific models were not reported, and they only rely on the aggregate models.

The most direct study related to this paper is by Tannura et al. (2008). They examined the relationship between weather, technology, and corn and soybean yields using Illinois, Indiana, and Iowa data from 1960 to 2006 in the U.S. Corn Belt. The estimated multiple regression models explained at least 94% and 89% of the variation in corn and soybean yields for each state, respectively.

The main framework of the multiple regression models is based on notable studies of Thompson discussed above. The key variables include precipitation, temperature, and technology. In terms of rainfall, there are pre-season precipitation (defined as the sum of precipitation over September of previous crop year through April of current crop year), May through August precipitation, and quadratic form of June through August precipitation. The temperature variables include May through August temperature. Finally, the annual time trend is presented as a proxy for technology advancement.

The estimated results showed that in the corn model, the pre-season precipitation was only significant in Iowa, and for each additional inch of rainfall the corn yield was expected to increase by one bushel per acre. May precipitation was significant in all three states, and the coefficients were negative. The explanation is that the wet weather in May would delay the planting progress and slow growth. June through August precipitation and their corresponding quadratic forms were mostly significant with July precipitation contributes the most to corn yields than any other month, and their sign of coefficients were as expected. Either too little or too much rainfall will negatively affect the corn yield, but the results showed that unfavorably dry weather would reduce yields more than favorably wet weather would increase yields. May and June temperatures had insignificant impact in all states and the magnitude were close to zero, although the sign of the coefficients turned out to be different in three states. The impact of July and August temperatures were significant in all states, and the low temperatures are more favorable to corn yields. The trend variable was highly significant in all three states and coefficients were close to 2, meaning technology by itself contributes to corn yield by 2 bushels per acre per year.

The modified soybean models did not work as efficiently as the corn models. Many variables were not statistically significant. The pre-season precipitation is only significant in Iowa, where one additional each of rainfall would increase soybean yields by 0.3 bushels per acre on average. May precipitation had negative and significant impact on soybean yields for all states, which is consistent with the case of corn. The impact of July precipitation was highly significant in Indiana, and July and August precipitations were highly significant in Iowa, but none of June through August precipitations was significant in Illinois. Speaking of magnitude, similar to corn, the damaging effect of dry weather is much more harmful than the wet weather is helpful. None of May through July temperatures was statistically significant in all three states, but May and June temperatures generally contribute positively to the soybean yield, and July and August temperatures had negative impact. The only variable that was highly significant in all three states is annual time trend, and it was estimated that technology would contribute about 0.5 bushels increase per acre per year.

Yield forecast evaluation was performed. Monthly out-of-sample crop yield forecasts from 1980 through 2006 were produced to compare with out-of-sample trend yield forecasts in June and July and with USDA Crop Production forecasts in August, September, and October. In some years, the corn and soybean yield forecast errors were even greater than one standard deviation, which ranged from about 15 to 19 bushels per acre and 5 to 9 bushels per acre, respectively. Forecasts from modified Thompson model were similar to the trend yield forecasts with soybean yield forecasts more accurate than corn yield forecasts. Modified Diebold-Mariano tests show that USDA forecasts were statistically more accurate than multiple regression model in both corn and soybean cases.

Irwin et al. (2008) modified their own model by adding another planting progress variable. They noticed two important aspects of planting date to yields. First of all, early planting is thought to contribute to the increasing overall trend in yields. They found corn and soybean planting start and complete dates of 2005 were about two weeks earlier than that of 1965. The second aspect is rooted in agronomic research. Several studies have shown that late planting will generally lead to lower yields than timely planting (Pecinovksy and Benson, 2001; Nafziger, 2008; Nielsen, 2008). For instance, in central Illinois, average corn yields decline at an accelerating rate for planting later than early May.

With the new late planting variable, Irwin et al. (2009a, 2009b) revised the multiple regression model for corn and soybean. May precipitation was dropped from both models to avoid multicollinearity problem brought up by implement of planting progress variable. April precipitation and its quadratic form entered the corn model, and consequently the duration of pre-season precipitation was defined as September of previous crop year through March of current crop year. Moreover, the quadratic form of pre-season precipitation was included in both models, and May and June temperature variables were dropped. New models were estimated with Illinois, Indiana, and Iowa data from 1960 to 2008, and explained at least 95% and 88% of the variation in corn and soybean yields for each state, respectively. The impact of planting progress variable on crop yields were significant and negative as expected, and other variables performed consistently with previous versions.

2.4 Summary

This chapter provides a review of agronomical literatures regarding plant physiology and some of important papers throughout the history of crop yield and weather studies in the U.S.

Corn Belt area. Agronomy theories suggest that the effect of temperature is different at different stages of crop development. For both corn and soybeans, there is a key phase, either silking or flowering. Before the key phase, high temperature boosts plant growth, which consequently contributes to high potential yield. Nonetheless, after corn silks or soybean flowers, high heat actually hastens crop development and shortens the grain or seed filling duration, thus reducing crop yields.

Many empirical studies resort to multiple regression model to explain the impact of weather on crop yield. The most recent one developed by Irwin et al. (2008, 2009a, 2009b) adopted Thompson's framework and brought in a planting progress variable, but the model was still based on monthly average weather variables using state level data. Non-linear model using growing degree days as a temperature variable attempted to reveal the extreme heat effect veiled by average temperature (Roberts and Schlenker, 2010, 2011; Schlenker and Roberts, 2006, 2008, 2009). Another study by Yu and Babcock (2010) created a new variable – drought index, which integrates both temperature and precipitation information. However, since the effect of heat and moisture was confounded together, the interpretation became troublesome. In sum, different approaches have been exercised in previous studies, but there is no one model that is superior to other models. The crop weather relations model is still an interesting topic requiring further development.

3. DATA

3.1 Introduction

As an empirical study focusing on estimation and evaluation of statistical models rather than crop simulation models, the key variables include corn and soybean yields, planting progress, precipitation, and temperature. This chapter introduces the source of each variable and explains how the variables are derived. In addition, descriptive analysis provides an overview of the data sets.

3.2 Yield, Weather, and Planting Progress Data

In the previous study by Tannura et al. (2008), crop-weather relations were analyzed in Illinois, Indiana, and Iowa, because these three states represented 43% to 45% of U.S. corn and soybean production from 2000 to 2006. In this study, the sample area extends to all major crop reporting districts (CRDs) in the U.S. Corn Belt area.

The major CRDs for corn and soybean production are derived in the following way. First, the percentage of CRD crop production to the U.S. total productions is calculated for every 5 years from 1960. If the production of a CRD is always greater than 0.25% of the U.S. total production, then it will be included in the sampling pool. However, there are some qualified CRDs that are far away from other CRDs and therefore become ‘islands.’ To keep the study in a certain closed area and avoid issues in spatial analysis, these CRDs are excluded. In addition, a few CRDs are added to the pool to fill the ‘hole.’ The resulting corn and soybean major production areas include 54 and 48 CRDs, respectively.

As pointed out by Tannura et al. (2008), a significant increase of fertilizer use began around 1960 (Thompson 1969, 1975, 1985, 1986, 1988; Garcia et al. 1987). To avoid the effect of nitrogen fertilizer being mixed with other weather variables, the starting date was specified at 1960.

Yield data were collected from USDA Quick Stats inquiry on the CRD level. Since for most CRDs in the sampling pool, crops are not irrigated, yield data derived from all practices were collected and used as the dependent variable in the analysis. However, in Nebraska as much as 90% of corn was irrigated in some years for some CRDs as Table 1 shows, so 3 heavily irrigated CRDs, Nebraska Central, Southwest, and South, were dropped from the corn sampling pool and the non-irrigated corn yield was used for another 3 less irrigated CRDs, namely Nebraska Northeast, East, and Southeast. In the case of soybeans, only a small portion of productions in Nebraska East district was irrigated, so no special treatment was needed. In addition, since soybean yield data of Kentucky was not available until 1972, two CRDs from that state were dropped.

Monthly average precipitation and monthly average temperature data at the CRD level comprise the key weather variables used in this analysis. They were collected through the National Climatic Data Center (NCDC) from the U.S. Temperature-Precipitation-Drought Index (DSI-9640) database. The series covered divisional monthly average precipitation, temperature, heating degree days, cooling degree days, and four drought indices from 1895 through the latest month available. Specifically, the time bias corrected data have been adjusted to rectify the bias induced by different recording time among observers (NCDC, 2002). The divisional monthly averages were calculated based on equally weighed station statistics within a climatic division, but the climatic divisions do not always coincide with USDA crop reporting districts. Thus, a few modifications were made. For instance, CRD 10 and CRD 20 in Missouri were merged as CRD 12 to form a better match to the climatic division. Resulting final maps include 49 original CRDs and one revised CRD for corn models (Figure 1), and 38 original CRDs and 4 revised CRDs for soybean models (Figure 2). Table 2 shows the percentage of sample CRDs' corn and soybean production to the U.S. total production from 1960 to 2010. The corn and soybean productions on average represent about 73% and 67% of total U.S. production.

Another way to capture the effect of heat on crop growth is introduced by the concept of growing degree days, which is usually measured as truncated degrees between upper and lower limits for a given period of time. For example, if the baseline was set at 10˚C, a day of 20˚C will contribute 10 growing degree days, and a day of 25˚C will contribute to 15 growing degree days. In a recent study by Roberts and Schlenker (2010), the moderate temperatures were defined as the total degree days between 10 and 29˚C for corn and between 10 and 30˚C for soybeans during the growing season, and extreme heat were defined as the total degree days above 29˚C and 30˚C during the growing season for corn and soybeans, respectively. Degree days data were provided by Michael Roberts from North Carolina State University and Wolfram Schlenker from Columbia University at the county level. To be consistent with yield and other weather variables, the county level data were transformed to the CRD level. Specifically, each county was matched to its corresponding CRD, and the degree days at the CRD level were calculated as a weighted average on harvested acreage on yearly basis.

Planting progress data were available only at the state level and were collected from USDA Quick Stats inquiry and *Weekly Weather and Crop Bulletin*. Since Quick Stats service only provides planting progress data dated back to 1979, late planting data from 1960 to 1978 were collected manually from *Weekly Weather and Crop Bulletin* published by U.S. Department of Commerce National Oceanic and Atmospheric Administration and USDA National Agricultural Statistics Service. The definition of late planting was adopted from Irwin et al. (2009a, 2009b). Based on agronomic recommendations from *Illinois Agronomy Handbooks*, planting progress variable was defined as the percentage of corn planted after May 30th from 1960 to 1985 and after May 20th from 1986 to 2010, or the percentage of soybeans planted after June 10th from 1960 to 1985 and after May 30th from 1986 to 2010. As explained by Irwin et al. (2008), late planting generally reduces crop yields. It is expected that the sign of coefficient of planting progress would be negative.

3.3 Descriptive Analysis

Before further analysis of crop weather relations using statistical models, descriptive statistics are summarized in this section. Since the sample sizes of the corn and soybean models are different, the summary statistics were calculated separately as shown in Table 3 and Table 4. Yield, late planting, and monthly weather data that were collected through public sources were summarized from 1960 to 2010. Growing degree days data provided by Michael Roberts from North Carolina State University and Wolfram Schlenker from Columbia University were only available up to 2005. In addition, to be consistent in data source, total precipitation data of March through August provided in the same data set were also summarized from 1960 to 2005, as they were used in modified Roberts and Schlenker model with growing degree days data.

3.3.1 Yields and Planting Progress

Corn and soybean yields have shown an increasing trend as displayed in Figure 3. The average corn yields in the sample CRDs increased at a rate of 1.82 bushels per acre per year, and the average soybean yields in the sample CRDs increased about 0.44 bushels per acre per year.

The average corn yield was 109.55 bushels per acre, and corn yields ranged from as low as 26.5 bushels per acre in East Central district of South Dakota in 1976 to 195 bushels per acre in Northwest district of Iowa in 2009. The average soybean yield was 35.43 bushels per acre, with a range of 46.1 bushels per acre. The lowest yield of 10.9 bushels per acre occurred in 1976 at West Central district of Minnesota, and the highest yield of 57 bushels per acre happened in 2010 at Central district of Illinois.

The late planting variable measures the late progress in crop planting. Since it is only available at state level, the statistics were summarized on state level observations. On average, about 19 percent and 30 percent of corn and soybeans were planted after the specified date, respectively. The biggest delay in corn planting happened in Ohio in 1996 due to rainy conditions. In case of soybeans, as much as 93.71 percent was planted after May 30th in 1991 in Tennessee because of heavy rainfall.

3.3.2 Temperature

Monthly average temperature from April through August was summarized for major corn production CRDs. The average temperature increased steadily until the largest value of 73.67°F was reached in July, and then August average temperature went down a little bit at 71.58°F. Pre-silking temperature defined as the average temperature of April, May, and June, and post-silking temperature defined as the average temperature of July and August reflect the overall temperature condition in these two seasons. For all temperature variables, the ranges were around 20°F except for April, and the standard deviations ranged from 2.83°F to 4.55°F, indicating normal weather conditions.

For major soybean production CRDs, monthly average temperature from May through August was displayed in Table 4. The similar pattern to corn sample was also found here with the maximum average temperature of 74.24°F occurred in July. Pre-flowering and post-flowering season temperatures were calculated as the average temperature of May through June, and July through August, respectively. The standard deviations were about 3°F and ranges were around 22°F. Both did not indicate unusual weather conditions.

Besides monthly average temperature, growing degree days data from 1960 through 2005 were also used as temperature variables. The average moderate heat in both corn and soybean samples was about 1400 growing degree days, which translate to about 18˚C or 64°F. Notice the average temperature of some days might be lower than 10˚C or higher than 29˚C or 30˚C, this translation is not a representation of average temperature during the growing season. The average extreme heat in corn sample was 26.48 growing degree days with a standard deviation of 20.28 growing degree days, and in soybean sample was 18.41 growing degree days with a standard deviation of 16.86 growing degree days, indicating a wide spread of observations, which was also verified by their wide ranges.

3.3.3 Precipitation

Pre-season precipitation was defined as the total precipitation over September of previous crop year through March or April of the current crop year in corn or soybean model, respectively. Naturally, the average pre-season precipitation of soybean CRDs was higher than that of corn CRDs. Interestingly, although the maximum precipitation was found in July with 16 inches or August with 16.66 inches, June appeared to be the wettest month with average precipitation of

4.23 inches and 4.28 inches in corn and soybean samples, respectively. Other months' average precipitations decreased gradually as they were further away from June.

Pre-silking and post-silking precipitation were defined as the total precipitation of April through June, and July through August in corn sample. Pre-flowering and post-flowering precipitation were defined as the total precipitation of May through June, and July through August in soybean sample. The ranges and standard deviations of these variables did not indicate any abnormal patterns.

The total precipitation during the growing season of March through August provided in the growing degree days data set was measured by centimeter. In both corn and soybeans CRDs, the total precipitation was around 55cm or 21.7 inches, which is consistent with the statistics of monthly data sets, assuming the monthly precipitation was around 3.5 to 4 inches.

3.4 Summary

This chapter presents the data sources, meaning of each variable, and descriptive analysis of the data sets. Yield and late planting data were collected from the USDA. An increasing trend exists in both corn and soybean yields. Planting progress varied from year to year and was greatly affected by weather conditions, since heavy rainfall will delay crop planting significantly. Monthly weather variables were obtained from NCDC, and some monthly variables were aggregated to seasonal variables in order to simplify the functional form in regression analysis. Pre- and post-silking temperature and pre- and post-flowering temperature were defined as the average temperature of the corresponding months, but aggregated precipitation variables were defined as the total precipitation of the matching months. The growing degree days data set was acquired from Michael Roberts from North Carolina State University and Wolfram Schlenker from Columbia University. The summary statistics were consistent with the results from the monthly weather variables data set.

4. MODELS

4.1 Introduction

This chapter introduces details of each statistical model that explains the relationship between crop yield and weather factors. Specifically, there are three main frameworks: modified Thompson model using monthly temperature, monthly precipitation, planting progress, and time trend variables; modified Robert and Schlenker model using growing degree days as a measurement of temperature and total precipitation; and panel geographically weighted regression (GWR) model that takes spatial dimension into consideration.

4.2 Modified Thompson Model

As a pioneer in crop yield and weather relation studies, Thompson developed a multiple curvilinear regression framework that utilizes monthly temperature, monthly precipitation, and time trend as independent variables. The functional form of corn and soybean models appeared in the 1985 and 1986 papers was

$$\begin{aligned}
y_t = {} & \beta_0 + \beta_1(\text{trend1})_t + \beta_2(\text{trend2})_t + \beta_3(\text{trend3})_t \\
& + \beta_4(\text{pre-season precipitation})_t + \beta_5(\text{pre-season precipitation})_t^2 \\
& + \beta_6(\text{July precipitation})_t + \beta_7(\text{July precipitation})_t^2 \\
& + \beta_8(\text{August precipitation})_t + \beta_9(\text{August precipitation})_t^2 \\
& + \beta_{10}(\text{June temperature})_t + \beta_{11}(\text{June temperature})_t^2 \\
& + \beta_{12}(\text{July temperature})_t + \beta_{13}(\text{July temperature})_t^2 \\
& + \beta_{14}(\text{August temperature})_t + \beta_{15}(\text{August temperature})_t^2 + \varepsilon_t
\end{aligned}$$

where y_t is the crop yield of a specific state in year t; trend1 through trend3 denote three time trends, specifically 1930 to 1959, 1960 to 1972, and 1973 to 1983, respectively; pre-season precipitation was defined as the sum of precipitation over September through June.

This model analyzed crop yield and weather relations in five major corn and soybean production states, i.e. Illinois, Indiana, Iowa, Missouri, and Ohio. By running the regression separately, the trend variables were estimated specifically for each state. Then, Thompson pooled the data together and ran one regression of crop yield on six weather variables to reveal the response of corn to weather factors. Thompson's model built up the foundation of regression analysis of crop weather studies. However, the analysis was constrained to specific states and the functional form did not consider the difference in corn and soybeans plant physiology, so further development taking those factors into consideration would be necessary.

Follow-up studies by Irwin et al. (2009a, 2009b) noticed the importance of planting progress in determining crop yields and revised Thompson's model. Corn and soybean models were specified separately rather than one single specification for both crops. Nonetheless, like previous Thompson models, revised models were still estimated on state level, i.e. Illinois, Indiana, and Iowa. Therefore, these models confined the estimation to individual state and lacked the ability to seize the overall effect from all observations. The corn and soybean models were specified as follows.

Corn model:

$$\begin{aligned}
y_t = {} & \beta_0 + \beta_1(\text{trend})_t + \beta_2(\text{late planting})_t \\
& + \beta_3(\text{pre-season precipitation})_t + \beta_4(\text{pre-season precipitation})_t^2 \\
& + \beta_5(\text{April precipitation})_t + \beta_6(\text{April precipitation})_t^2 \\
& + \beta_7(\text{June precipitation})_t + \beta_8(\text{June precipitation})_t^2 \\
& + \beta_9(\text{July precipitation})_t + \beta_{10}(\text{July precipitation})_t^2 \\
& + \beta_{11}(\text{August precipitation})_t + \beta_{12}(\text{August precipitation})_t^2 \\
& + \beta_{13}(\text{July temperature})_t + \beta_{14}(\text{August temperature})_t + \varepsilon_t
\end{aligned}$$

where late planting variable was defined as the percentage of corn planted after May 30th from 1960 to 1985 and after May 20th from 1986 to 2008; pre-season precipitation was defined as the total precipitation over September of previous crop year through March of current crop year.

Soybean model:

$$
\begin{aligned}
y_t = \beta_0 + \beta_1(\text{trend})_t + \beta_2(\text{late planting})_t \\
&+ \beta_3(\text{pre-season precipitation})_t + \beta_4(\text{pre-season precipitation})_t^2 \\
&+ \beta_5(\text{June precipitation})_t + \beta_6(\text{June precipitation})_t^2 \\
&+ \beta_7(\text{July precipitation})_t + \beta_8(\text{July precipitation})_t^2 \\
&+ \beta_9(\text{August precipitation})_t + \beta_{10}(\text{August precipitation})_t^2 \\
&+ \beta_{11}(\text{July temperature})_t + \beta_{12}(\text{August temperature})_t + \varepsilon_t
\end{aligned}
$$

where late planting variable was defined as the percentage of soybeans planted after June 10th from 1960 to 1985 and after May 30th from 1986 to 2008; pre-season precipitation was defined as the total precipitation over September of previous crop year through April of current crop year.

Since soybeans are generally planted later than corn, these two models reflect this observation in the definition of late planting and pre-season precipitation variables, as well as the functional form. April precipitation entered the corn model, but May precipitation was dropped from both models due to the multicollinearity problem raised by implement of planting progress variable. May and June temperature variables were also dropped because they were not statistically significant.

To enhance the overall explanatory power of the regression and better reflect the findings in agronomy literature, Thompson type models were modified to a new stage. The functional forms for corn and soybean model were specified as follows.

Corn model:

$$
\begin{aligned}
y_{it} = {} & \alpha_i + \beta_0 + \beta_1(\text{trend})_t + \beta_2(\text{late planting})_{jt} \\
& + \beta_3(\text{pre-season precipitation})_{it} + \beta_4(\text{pre-season precipitation})_{it}^2 \\
& + \beta_5(\text{April precipitation})_{it} + \beta_6(\text{April precipitation})_{it}^2 \\
& + \beta_7(\text{May precipitation})_{it} + \beta_8(\text{May precipitation})_{it}^2 \\
& + \beta_9(\text{June precipitation})_{it} + \beta_{10}(\text{June precipitation})_{it}^2 \\
& + \beta_{11}(\text{July precipitation})_{it} + \beta_{12}(\text{July precipitation})_{it}^2 \\
& + \beta_{13}(\text{August precipitation})_{it} + \beta_{14}(\text{August precipitation})_{it}^2 \\
& + \beta_{15}(\text{pre-silking temperature})_{it} + \beta_{16}(\text{post-silking temperature})_{it} + \varepsilon_{jt}
\end{aligned}
$$

where y_{it} is the corn yield in CRD i and year t; late planting and pre-season precipitation were defined the same as above corn model in Irwin et al. (2009a), but notice the planting progress data were only available on state level; pre- and post-silking temperature were defined as the average temperature of April through June, and July through August, respectively; the error term ε_{jt} was clustered by state.

Soybean model:

$$
\begin{aligned}
y_{it} = {} & \alpha_i + \beta_0 + \beta_1(\text{trend})_t + \beta_2(\text{late planting})_{jt} \\
& + \beta_3(\text{pre-season precipitation})_{it} + \beta_4(\text{pre-season precipitation})_{it}^2 \\
& + \beta_5(\text{pre-flowering precipitation})_{it} + \beta_6(\text{pre-flowering precipitation})_{it}^2 \\
& + \beta_7(\text{post-flowering precipitation})_{it} + \beta_8(\text{post-flowering precipitation})_{it}^2 \\
& + \beta_9(\text{pre-flowering temperature})_{it} + \beta_{10}(\text{post-flowering temperature})_{it} + \varepsilon_{jt}
\end{aligned}
$$

where y_{it} is the soybean yield in CRD i and year t; late planting and pre-season precipitation were defined the same as above soybean model in Irwin et al. (2009b), again notice the planting progress data were only available on state level; pre- and post-flowering temperature were defined as the average temperature of May through June, and July through August, respectively; and the error term ε_{jt} was clustered by state.

Two major modifications of Thompson type model were made in this thesis. First, for both corn and soybean, the sample size was expanded to all the major CRDs in the U.S. and a fixed effects model was employed rather than individual state level model. Along with increase of observation numbers, one statistical issue emerged, i.e. standard errors might be biased downward. Thus, a robust standard error clustered by state was calculated to ensure that the significance level of independent variables was not overstated.

A second novel change to the Thompson model was the introduction of the concept from plant physiology and phenology. As the agronomy literature indicated in section 2.2, the key boundaries that classify the role of temperature in plant development and seed filling are silking and flowering for corn and soybeans, respectively. Before silking or flowering period is the vegetative stage, when high or moderate heat helps plants grow their equipment to seed; however, after silking or flowering period, the crop enters into reproductive stage, in which high temperatures accelerate crop development and shorten grain filling duration, resulting in small seeds, and therefore low yield.

4.3 Modified Robert and Schlenker Model

Besides the most common way to measure heat using average temperature over a season or a month, Robert and Schlenker employed growing degree days, arguing that the non-linear effect of extreme heat was diluted when temperatures were averaged. The fine-scale growing degree days data set was constructed based on estimation of time that a crop is exposed to each 1 degree Celsius interval, and therefore may provide a better way to identify the impact of extreme heat and non-linear effect of temperature.

To build the regression model, Robert and Schlenker made an important assumption. They postulated that the temperature effect on plant growth is cumulative over time and yield is proportional to total growth, which implies that the temperature effect on yield is additive and substitutable over time. This assumption is implicitly justified by a statistically significant relationship between the cumulative distribution of temperatures and yields. They argued that random paired temperatures should not provide a clear identification if time separability and substitutability were not appropriate. The regression model showing the relationship between log yield y_{it} and yield growth $g(h)$ which non-linearly depends on heat h in county i and year t is

$$\log(y_{it}) = \int_{\underline{h}}^{\overline{h}} g(h)\phi_{it}(h)dh + \mathbf{z}_{it}\boldsymbol{\delta} + \alpha_i + \varepsilon_{it}$$

where $\phi_{it}(h)$ is the time distribution of heat over growing season in county i and year t. The growing season was defined as months March through August for both corn and soybeans in this model, $\mathbf{z}_{it}$ denotes other independent variables including total precipitation and its quadratic term, a time trend and its quadratic form for each state that captures technological improvement, α_i is a county fixed effect intercept that controls for heterogeneity, and ε_{it} is the error term which allows for spatial correlation.

Robert and Schlenker specified three forms of $g(h)$, which are rather mathematically complex in their early studies. Nonetheless, a simple model that commits to the principle of parsimony was introduced in their 2010 publication (Robert and Schlenker, 2010). The baseline model was specified as

$$\log(y_{it}) = \alpha_i + \beta_1 h_{it} + \beta_2 m_{it} + \beta_3 p_{it} + \beta_4 p_{it}^2 + \beta_5 t_{jt} + \beta_6 t_{jt}^2 + \varepsilon_{it}$$

where y_{it} are corn and soybean yields in county *i* and year *t*; α_i is a county fixed effect; h_{it} denotes extreme heat (degree days above 29˚C and 30˚C for corn and soybeans, respectively); m_{it} denotes moderate temperatures (degree days 10-29˚C and 10-30˚C for corn and soybeans, respectively); p_{it} is the total precipitation over the growing season; t_{jt} is state-specific time trend; and ε_{it} is the error term.

This simple fixed effect model provides a straightforward view of the factors that affect the crop yield. Two temperature variables, extreme heat and moderate heat, clearly play an indispensible role in the model. While the moderate heat generally contributes positively to crop yields, extreme heat is expected to affect the yield in a negative way. In addition, since heat was measured using growing degree days rather than average temperature, the non-linear temperature effect will be captured. Two precipitation terms identify the quadratic effect of moisture. However, since the growing degree days data set only record total precipitation, to keep the data source consistent, monthly precipitation was not employed in this model. Finally, the state-specific time trend and its quadratic term capture technology advancement.

To be consistent with other models in scale and make comparable evaluations of predictive powers between models, the modified Robert and Schlenker model is specified on CRD level:

$$y_{it} = \alpha_i + \beta_0 + \beta_1 h_{it} + \beta_2 m_{it} + \beta_3 p_{it} + \beta_4 p_{it}^2 + \beta_5 t_t + \varepsilon_{jt}$$

where the notations are basically the same with the baseline Robert and Schlenker model except for that *i* denotes CRD rather than county, and ε_{jt} is the error term clustered by state. There are two major revisions in the modified model. First, the time trend variable does not differentiate between states and the quadratic time trend is no longer included in the model. Studies by Irwin et al. (2009a, 2009b) analyzed determinants of crop yield with a trend variable in three major corn and soybean production states Illinois, Indiana, and Iowa. The results showed that the maximum difference of trend yield estimates between these three states was 0.28 bushels for corn, and 0.07 bushels for soybeans. Thus, a single trend variable for all observations as a proxy for technology is reasonable. Second, the quadratic trend term was dropped because it did not improve the model performance. By comparing the estimation results with and without the quadratic trend term, the R-squared value actually decreased and the coefficients of other independent variables had tiny changes, so the simpler model was adopted to increase degrees of freedom.

As explained in modified Thompson's model, standard errors tend to be underestimated when irrelevant observations entered the estimation. Therefore, robust standard errors were obtained by clustering on state level, meaning only CRDs in the same state were considered when calculating standard errors for a specific CRD.

Comparing the modified Robert and Schlenker model with the modified Thompson model, one of the key differences lies on the treatment of the temperature variables. The modified Thompson model takes plant physiology into consideration. The temperature effect is believed to be different before and after silking or flowering phase for corn or soybeans. Nonetheless, in the modified Robert and Schlenker model, a basic assumption is that a particular temperature will have the same effect on crop yield throughout all the growing season no matter it is in May or in August. Although the assumption may not be consistent with agronomy theories, it was verified from statistical evidence. While both models are plausible, further comparison on model performance will be left on their predictive power assessment.

4.4 Panel Geographically Weighted Regression Model

The previous two sections introduced two fixed effect models with different explanatory variables. In essence, by placing a fixed effect term in the specification these two models assume that independent variable coefficients are all the same across sample CRDs, but only intercepts differ. While keeping the statistical model simple, this assumption allows spatial heterogeneity to a very low degree. To examine spatial heterogeneity issue more carefully, a recently developed technique, geographically weighted regression (GWR), was adopted.

As the name geographically weighted regression itself implies, GWR is basically a weighted regression scheme using weights generated from geographical information. Fotheringham et al. (2002) illustrated the GWR model in detail. Consider an ordinary linear regression

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i .$$

To incorporate geographical information, it is necessary to identify each observation's coordinate. Mathematically, the basic linear geographically weighted regression can be expressed as

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i$$

where (u_i, v_i) denotes the geographic coordinates of observation i in space. Then $\beta_k(u_i, v_i)$ represents the local parameter at location (u_i, v_i), which is a realization of the continuous function $\beta_k(u, v)$ at point i. If all the coefficients are spatially invariant, the above equation can be simplified to the ordinary linear regression model, i.e. the ordinary linear regression can be viewed as a special case of GWR.

In practice, to run a GWR, the coordinates are used to calculate distances between each pair of observations, and then all the distances are transformed into a weight matrix. Essentially, the spatial weight matrix is constructed so that observations closer to the regression point will be given more weight in the local regression than observations farther away.

To construct a weight matrix, it is necessary to define a spatial kernel and its bandwidth. Obviously, a spatial kernel should reflect the property that the spatial weight w_{ij} decreases as the distance d_{ij} between the regression point i and the data point j increases as Figure 4 shows. There are two kinds of spatial kernels, namely fixed spatial kernels and adaptive spatial kernels. In a fixed spatial kernel, observation points with the same distance to the regression point will be given the same weight no matter how dense or sparse the data are (Figure 5), whereas an adaptive spatial kernel should adapt itself to the density of the data so that the case of too few points in a local regression can be avoided (Figure 6). Several common spatial kernels have been studied. For example, Gaussian kernel,

$$w_{ij} = \exp[-\frac{1}{2}(d_{ij} / b)^2]$$

bi-square kernel,

$$\begin{aligned} w_{ij} &= [1-(d_{ij} / b)^2]^2 \text{ if } d_{ij} < b \\ &= 0 \text{ otherwise} \end{aligned}$$

and inverse distance kernel,

$$\begin{aligned} w_{ij} &= 1/ d_{ij} \text{ if } d_{ij} < b \\ &= 0 \text{ otherwise} \end{aligned}$$

where w_{ij} is the spatial weight, d_{ij} is the distance between the regression point i and the data point j, and b is the bandwidth. Among the above three kernels, the first two are fixed spatial kernels and the last one is an adaptive spatial kernel. The choice of bandwidth should be careful. As explained by Fotheringham et al. (2002), this distance-decaying measure determines the extent to which the local regression results are smoothed. Smaller bandwidth usually results in rougher and steeper surface than larger bandwidth.

The discussion so far only applies to cross sectional data, but not to panel data with both time and space dimensions. To estimate all parameters for all CRDs at the same time, a panel geographically weighted regression was created. Before illustrating the data structure of panel GWR, Table 5 shows an example of three year data with three regions A, B, and C. Using the weight pattern of this data set, the regression results will only give appropriate coefficient estimate for region A. To incorporate all information and estimate coefficients for all regions together, a large data set was constructed as shown in Table 6. Essentially, the coefficients for each region will be estimated with its own independent variables, for instance, region B's coefficients corresponds to the estimate of regressor vector X_B. With such data structure, all regions are treated properly.

Based on Akaike information criterion (AIC) and Bayesian information criterion (BIC) of models with different kernels, the best kernel and bandwidth is bi-square kernel with the minimum bandwidth that ensures each CRD should have at least one neighbor. The acquisition of a fixed spatial kernel is reasonable. Since all CRDs are about similar size and there is no isolated CRD, the centroids should by itself scatter relatively evenly on the map without resorting to adjustment for density issues. A small bandwidth was obtained largely due to the fact that the number of non-zero weight points will increase as the bandwidth became larger, and therefore in a panel GWR scheme the number of parameters will go up, which enlarges AIC and BIC scores. As a result, evaluation of panel GWR will be performed with bi-square kernel and minimum bandwidth.

4.5 Summary

Three statistical models are illustrated in this chapter. Each model has its own advantages and also its shortcomings, and there is no one particular model that is superior to all other models. While keeping the minimal number of variables, the model also loses the potential to draw conclusions from other useful information. It’s a tradeoff between parsimony principle and amount of information. Therefore, all three models are exercised to forecast crop yields. In the next chapter, estimation results will be presented.

5. EMPIRICAL RESULTS

5.1 Introduction

This chapter examines results of each regression model. All models were estimated with the most recent data available. Specifically, modified Thompson models and panel GWR models were estimated with data from 1960 through 2010, and modified Robert and Schlenker models were evaluated with data from 1960 through 2005. Model performance statistics and coefficient estimates will be reviewed and interpreted individually. A comparable assessment of model performance in terms of predictive power will be given in the next chapter.

5.2 Results of Modified Thompson Model

Results of modified Thompson corn and soybean models are displayed in Table 7 and Table 8. Both models performed decently, with R-squared of 88% and 84% for corn and soybean model, respectively, meaning 88% and 84% of variation of yield can be explained by these two models. Compared with results of two recent studies by Irwin et al. (2009a, 2009b) with similar specifications, the R-squared values of fixed effect models were not as high as those of state level models, which were higher than 95% for corn models and around 90% for soybean models. The standard errors of corn and soybean models are about 12 bushels per acre and 3.5 bushels per acre, respectively, which reflect model's R-squared statistic properly. One possible explanation is that with more detailed data on CRD level, there might be more variations need to be explained than data on state level, even though the functional form of the regression improved. F-stats of both models were highly significant at the 1% level, indicating the data fit the regression model very well.

In the modified corn model, the trend variable was highly significant with a coefficient estimate of 1.86 bushels per acre, meaning on average the technology change contributes 1.9 bushels per acre increase to corn yield every year. The coefficient of late planting variable was estimated at -0.25 bushels per acre with 1% significance level. According to the definition of late planting variable, the interpretation is that for every 1% of corn planted after May 30th from 1960 to 1985 or after May 20th from 1986 to 2010, the corn yield will be reduced by 0.25 bushels per acre. This number has a notable meaning. For example, in 2009, 62% of corn was planted later than May 20th in Illinois, which translates to 15.5 bushels per acre decrease in corn yield. The effect is enormous.

All precipitation variables were highly significant and the sign of coefficients indicated concave effects to the corn yield as expected. Figure 7 shows the response of corn yields to precipitation variables. Notice the quadratic form in the specification, the marginal effect varies at different values, and since the estimates showed that the squared term was negative, the marginal effect decreases as the amount of rainfall increases. Pre-season precipitation measures the amount of moisture in soil before planting that contributes to the corn yield. The marginal effect of pre-season precipitation at the mean was 0.22 bushels per acre, meaning on average one additional inch of pre-season rainfall above the mean value will increase the corn yield by 0.22 bushels per acre. Keeping other variables constant, the maximum potential corn yield was reached with 19.74 inches of pre-season rainfall, which was 3.5 inches higher than the current average value. The marginal effects of April and May precipitation at the mean were 0.88 and 0.28 bushels per acre, respectively. Given the mean values and the standard deviations of these two months' precipitation, their impact on yield was relatively small. On the other hand, the marginal effects of June and July precipitation at the mean were 1.37 and 3.94 bushels per acre, as Panel D and Panel E of Figure 7 shows. The marginal effect of August precipitation on corn yields was 0.94 bushels per acre at the mean value of 3.68 inches, which was about 3 inches less than the optimal level.

Although the impact of pre-silking temperature was not statistically significant, the sign of two temperature variables confirmed the fact that the impact of heat was different in determining corn yield at different phases of crop development. The response of corn yields to temperature variables was presented in Figure 8. One degree Fahrenheit increase in pre-silking temperature was estimated to raise the corn yield by 0.21 bushels per acre, while one degree Fahrenheit increase in post-silking temperature would reduce the corn yield by 3.4 bushels per acre. Thus, high heat during the post-silking period was expected to harm corn yields more severely than the benefit effect from the pre-silking temperature.

In the modified soybean model, the impact of technology trend was significant at the 1% level and was estimated to bring 0.45 bushels per acre increase to soybean yields every year, which is comparable to the trend only model estimate as panel B of Figure 3 indicates. The coefficient of late planting variable was -0.048 bushels per acre with 1% significance level, meaning for every 1% of soybeans planted after June 10th from 1960 to 1985 or after May 30th from 1986 to 2010, the soybean yield was expected to decrease by about 0.05 bushels per acre.

The impact of pre-season precipitation on soybean yields was significant at the 10% level, and the marginal effect at the mean was 0.03 bushels per acre. That means at the average pre-season precipitation value of 21 inches, one additional inch of rainfall would increase the soybean yield by 0.03 bushels per acre. Both pre-flowering and post-flowering precipitation were significant at the 1% level. The marginal effect of pre-flowering precipitation at the mean was

0.2 bushels per acre, and extra 1.4 inches of rainfall will be needed to reach the maximum soybean yields as Panel B of Figure 9 shows. The impact of post-flowering precipitation was the greatest among these three precipitation variables. It was calculated that its marginal effect on soybean yields was 0.86 bushels per acre at the current average value of 7.77 inches. Keeping other variables constant, the maximum soybean yield will be reached with 11.3 inches of post-flowering rainfall.

Estimation of pre-flowering temperature was statistically significant at the 1% level, and the coefficient indicated that on average one additional degree Fahrenheit increase in pre-flowering temperature will raise the soybean yield by 0.24 bushels per acre. High post-flowering temperature contributed negatively to the soybean yield as shown in Panel B of Figure 10. One degree Fahrenheit increase in post-flowering temperature will reduce the soybean yield by 0.38 bushels per acre. Similar to results from modified corn model, the harmful impact of high temperature on soybean yields during the post-flowering period was greater than the positive impact during the pre-flowering period.

In summary, results from modified soybean models were consistent with findings in modified corn model, i.e. the impact of both precipitation and heat during the reproductive stage was greater than their impact during the vegetative stage. This makes sense because the most essential process that directly relates to crop yields – seed filling – occurs in the reproductive stage. Thus, one important implication of modified Thompson model is that severe weather conditions during July and August would affect crop yields much greater in magnitude than unbeneficial weathers during other periods.

5.3 Results of Modified Robert and Schlenker Model

The regression results of modified Robert and Schlenker model for both corn and soybeans are presented in Table 9. The overall performance of both models in terms of R-squared was not as good as modified Thompson models. The modified Robert and Schlenker corn and soybean models explained about 84.5% and 78% of total variations in crop yield, respectively, as compared to 88% and 84% variations being explained in modified Thompson models. The standard errors of corn and soybean models were 12.31 bushels and 3.81 bushels, respectively, which were slightly higher than the results of modified Thompson models.

For the corn model, all independent variables are statistically significant at the 1% level except the intercept. The coefficient of trend variable signifies that on average the technology advancement will bring up corn yield by 1.7 bushels per acre every year. This number is very close to the results from modified Thompson model, which is about 1.9 bushels per acre.

The impact of total precipitation over the growing season on corn yields was presented in Figure 11. The marginal effect at the mean was -0.16 bushels per acre. Notice the unit of precipitation was centimeter, so the number translates to that at the average value of 54.81cm or 21.6 inches, one additional centimeter of total rainfall will reduce the corn yield by 0.16 bushels per acre, or one additional inch of total rainfall will reduce the corn yield by approximately 0.4 bushels per acre. One interesting observation is that it is the only model among all corn and soybean models implying there was too much rainfall. Figure 11 shows that the average total precipitation exceeded the optimal value that maximizes crop yields by 4cm or 1.6 inches. To explain this result, a review of model specification is helpful. In modified Thompson model, growing season for corn was defined as April through August, whereas in modified Robert and

Schlenker model, the total precipitation was calculated as the sum of rainfall over March through August. Thus, inclusion of March in the growing season might overemphasize the precipitation effect.

The directions of moderate and extreme temperatures effect were as expected, but the magnitudes of their effect differed a lot. As Figure 12 shows, while a 10 moderate growing degree days increase was expected to raise corn yields by 0.3 bushels per acre, a 10 extreme growing degree days increase would reduce corn yields by 7.5 bushels per acre. In other words, the extreme hot weather would reduce the yield much more than the beneficial moderate heat would increase the yield.

In the modified soybean model, again all the independent variables are statistically significant at the 1% level expect the constant term. The technology trend was estimated at 0.42 bushels per acre, meaning on average technology improvements enhances soybean yields by 0.42 bushels per acre every year, which is comparable with the trend estimate of 0.45 bushels per acre from modified Thompson model.

The marginal effect of total precipitation on soybean yields at the mean was 0.07 bushels per acre, indicating at the mean value of 56.42cm or 22.2 inches, one additional centimeter of total rainfall will increase the soybean yield by 0.07 bushels per acre, or one additional inch of total rainfall will raise the soybean yield by 0.18 bushels per acre. Figure 13 shows that the optimal total precipitation was 8.2cm or 3.2 inches higher than the current average precipitation level.

The effect of temperature variables were shown in Figure 14. The coefficients of moderate heat and extreme heat were 0.013 and -0.184, respectively. Like the results from modified corn model, the magnitude of the coefficients indicated the beneficial effect of moderate heat would not offset the adverse impact of extreme heat on soybean yields at the same level. Given the mean value and standard deviation of extreme heat as shown in Table 4, the impact of unfavorable hot weather could be big. If the extreme temperature goes up by 20 growing degree days, then the soybean yield would decrease by almost 4 bushels per acre.

To sum up, estimation results from modified Robert and Schlenker models explained yield variations from another prospect. Due to different definition of growing season, the average total precipitation turned out reducing corn yields, as compared to results from other models which indicated the optimal rainfall value that maximized crop yields should be above the average level. While the key assumption of modified Robert and Schlenker model did not differentiate the temperature effect during different plant stages, the results showed that the impact of extreme heat throughout the entire growing season was greater than the moderate temperature effect.

5.4 Results of Panel Geographically Weighted Regression Model

In the panel geographically weighted regression model, essentially, each CRD was estimated with its own coefficients. By plotting the coefficients on a map, one would hope to discover the geographical behavior pattern of each independent variable, and then aggregate similar CRDs to a group, which finally leads to region-specific models. Thus, to fully illustrate the features of coefficient estimates, results were presented in form of both tables and figures.

Table 10 and Table 11 display the results from panel GWR corn model. Generally speaking, the estimates were very close to the modified Thompson corn model. The coefficients of trend variable were all statistically significant at a 1% level and ranged from 1.51 to 2.17

bushels per acre. The average was 1.84 bushels per acre, almost identical to the estimate from modified Thompson corn model. Figure 15 showed that the technology change contributed highest to corn yields in the northwest region of Corn Belt area. The late planting variable was not significant in a couple of CRDs, but they all showed harmful effect on corn yields, and the average effect was -0.28 bushels per acre. Similar to the regional effect of technology advancement, the impact of late planting activities was greatest to the northwest, and least in southern Illinois and western Indiana as Figure 16 indicated.

Precipitation variables turned out to have mixed effect for different CRDs. The impact of pre-season precipitation was not statistically significant at a 10% level for half of the CRDs, and in southern Minnesota and southern Wisconsin the sign of coefficients indicated that the marginal effect of rainfall would increase as precipitation went up. A careful examination of panel GWR will help to explain this odd finding. In estimating coefficients for each CRD, only nearby districts that have positive geographical weights were used in the regression. As a result, the number of observations was small for a specific CRD, and the coefficients only reflect patterns in those CRDs. For instance, notice the estimated coefficients were identical for two CRDs, MN-40 and MN-50. This is because these two CRDs were far away from other districts and therefore identified each other as their only neighbor, so the estimation of both CRDs was in fact based on the same subset.

Now that it is clear the estimates for individual CRD may not necessarily reveal the general pattern of crop yields and weather relations, a review of average estimates might be more helpful. The mean value of marginal effect of pre-season precipitation at the mean was 0.33 bushels per acre, indicating on average each additional inch of rainfall above the current mean value would raise corn yields by 0.33 bushels per acre. Figure 17 showed that CRDs to the west side had the greatest pre-season precipitation marginal effect. April and May precipitations were not significant at a 10% level in majority of CRDs. The average marginal effect at the mean was 0.68 and 0.47 bushels per acre for April and May precipitations, respectively. On the other hand, June and July precipitation was mostly significant at a 1% level. Their average marginal effect at the mean was 1.61 and 4.23 bushels per acre, respectively, which were similar to the results from modified Thompson model. As for August precipitation, the marginal effect at the mean ranged from -0.76 to 3.15 bushels per acre, with an average of 1.3 bushels per acre, higher than the estimate from modified Thompson corn model. Figure 18 through Figure 22 plotted the marginal effect of all monthly precipitation variables, but none of them showed obvious clustering patterns.

Pre-silking temperature was estimated to contribute positively to the corn yield with a few exceptions. The average of coefficients was 0.52 bushels per acre, more than twice high of estimate from modified Thompson corn model. In addition, notice the range was from -1.79 bushels per acre in MI-90 to 1.88 bushels per acre in IL-10, there were wide variations in the impact of pre-silking temperature. Figure 23 did not show any clear pattern either. For post-silking temperature, all coefficients were significant at a 1% level except for district MI-90, and the negative sign was as expected. The average effect was -3.2 bushels per acre, indicating on average each degree Fahrenheit increase in post-silking temperature would reduce corn yields by 3.2 bushels per acre. By looking at Figure 24, it seems CRDs to the north were less affected than CRDs to the south. One possible reason is that northern districts are generally colder than southern regions, so the adverse impact of high post-silking temperature was less severe to the north.

Results of the panel GWR soybean model were presented in Table 12 and Table 13. Trend variable was significant at a 1% level for all 42 CRDs. The average value of coefficients was 0.45 bushels per acre, the same as results from modified Thompson soybean model. Figure 25 showed that the northwestern region had the highest trend while the southern part had the lowest. The late planting variable was significant in most CRDs and showed negative impact on all districts. On average, for each 1% soybeans planted after June 10^{th} from 1960 to 1985 or after May 30^{th} from 1986 to 2010, the soybean yield would decrease by 0.053 bushels per acre. This number was only 0.005 bushels per acre higher than the estimate from modified Thompson model. The plotted map did not indicate any noticeable grouping pattern.

The impact of pre-season precipitation was only significant in a few CRDs, and the sign of coefficients may not necessarily meet our expectations. The average marginal effect at the mean was 0.023 bushels per acre, slightly lower than 0.026 bushels per acre in modified Thompson soybean model. Pre-flowering precipitation aggregated rainfall during May and June, and the total effect was mostly significant. The marginal effect at the mean ranged from -0.067 to 0.58 bushels per acre, and the average was 0.24 bushels per acre, meaning on average each additional inch of pre-flowering rainfall above the current mean would bring up the soybean yield by 0.24 bushels per acre. Coefficients of post-flowering precipitation were highly significant at a 1% level in entire area, and all signs were as expected. The average marginal effect at the mean was 0.89 bushels per acre, higher than 0.86 bushels per acre in modified Thompson model. Figure 27 through Figure 29 suggested no clustering patterns for these precipitation variables.

The impact of pre-flowering temperature was not statistically significant at a 10% level for a few CRDs in states like Indiana and Minnesota, and two CRDs MN-40 and MN-50 had shown a harmful impact of pre-flowering heat on soybean yields. Nonetheless, on average one degree Fahrenheit increase in pre-flowering temperature was expected to raise the soybean yield by 0.3 bushels per acre. Figure 30 indicated that Missouri and southeastern Iowa benefitted the most from pre-flowering heat, whereas the impact was minimum and even negative in southern Minnesota. Post-flowering temperature showed consistent and significant adverse impacts through the entire sample of CRDs. The average impact on the soybean yield was -0.4 bushels per acre for each degree Fahrenheit, slightly higher than -0.38 bushels per acre of modified Thompson soybean model in magnitude. Figure 31 showed that the impact was greatest in southwestern Minnesota and northwestern Iowa.

In general, results from panel GWR corn and soybean models were consistent with findings in modified Thompson models. The average effect of independent variables was similar to those in aggregate fixed effects model, and the magnitude of coefficients was greater in reproductive stage than in vegetative stage. However, the plots of coefficients did not show stable clustering patterns for either crop. As a result, it is impossible to group CRDs and the panel GWR models were adopted as the final model.

5.5 Summary

Results from three different models were examined in this chapter. Modified Thompson models and panel GWR model yielded comparable results, which showed that impact of both precipitation and heat during the reproductive stage was greater than their impact during the vegetative stage. In addition, results indicated that the marginal effect of precipitation variables in modified Thompson models had not achieved the maximum at the current average value. On the other hand, modified Roberts and Schlenker models were limited due to incomplete data set only up to 2005 and their specification of temperature variable, which assumed uniformed impact during vegetative and reproductive stages. To evaluate model performance in terms of forecasting ability, three models' predictions of crop yields will be assessed in the next chapter.

6. A YIELD FORECASTING COMPETITION

6.1 Introduction

One of the important uses of a crop weather model is yield forecasting. An important publication that provides monthly predictions of crop production in the U.S. is *World Agricultural Supply and Demand Estimates* (WASDE) by USDA. It provides useful estimates of crop yield, production, stocks, price, and so forth for all agriculture related businesses. As an alternative, predictions of corn and soybean yields from all models will be studied in this chapter, and their predictive power will be evaluated based on several different criteria.

Two benchmarks for comparison were set as: 1) USDA September forecast and 2) trend-only yield forecast. USDA publishes its crop yield forecast in August, September, October, and November in *Crop Production* every year. The September forecast was picked to ensure that all predictions were based on the same availability of information, because weather variables were included in all regression models up through August. Trend-only forecasts were obtained with a regression model without any weather information but only time trend. It is the simplest model that corresponds to the argument that crop yield will increase steadily due to new technology.

6.2 Development of the Yield Prediction Competition

Previous studies have specified two methods to evaluate model prediction performance. Tannura et al. (2008) adopted the conventional recursive forecasting model. Forecasts were produced monthly at the beginning of June, July, August, September, and October from 1980 to 2006. Specifically, since the actual weather values were not available until the end of growing season, the forecast was obtained with both actual and expected weather variables. For instance, actual weather values from 1960 to May 1980 and average weather values of other variables from 1960 to 1979 were used to predict crop yield of June 1980. Then, in the next estimation period, actual weather values of June 1980 will replace expected values to predict crop yield of July 1980. By repeating this process at the end of each prediction month, forecasts will be updated with more accurate weather information. Essentially, in a recursive forecasting scheme, new observations will be added one at a time to make new predictions.

Schlenker and Roberts (2009) evaluated out-of-sample predictions using a resampling strategy similar to the idea of bootstrapping. They randomly picked 48 out of 56 years data, and ran a regression with the sub-sample to predict crop yields of the remaining 8 years. The argument was that because yields are spatially correlated within any given year, the sampling process should be based on whole years but not on observations. Otherwise, the spatial structure would be broken across different years. This process was repeated 1000 times to obtain an efficient estimation.

While the recursive method emphasizes models' ability to forecast yields in the future, the resampling strategy pays more attention to prediction accuracy without considering the time direction of forecasts. Both methods will be exercised in this thesis. Since forecasts from modified Roberts and Schlenker models were only available up to 2005, two separate competitions were held. Specifically, two sets of recursive forecasts statistics were calculated from 1988 through 2005 and from 1988 through 2010, respectively. As for the bootstrapping forecasts, 6 year sub-samples were drawn from the 1960 through 2005 data set, and 7 year sub-samples were drawn from the 1960 through 2010 data set. Both sub-samples were about 13 to 14% of the corresponding sampling pool.

Before delving into the discussion of evaluation standards, it is important to answer this question: which yield are we predicting? Naturally, since all models were estimated on the CRD level, one most direct answer would be crop yields of all major CRDs in the sampling pool. However, a more important value that people are interested in is the projection of total U.S. crop yield. In order to translate CRD level yields to the U.S. yield, the methodology used by Irwin et al. (2009a, 2009b) was adopted. In the first step, the harvested acreage-weighted average of all CRDs' predicted yield was calculated. Then, the ratio of harvested acreage-weighted average of all CRDs' actual yields to the U.S. actual yield over the past 10 years was also computed. In the final step, the U.S. yield forecast was obtained by adjusting the area weighted average of predicted yield using the past 10 years' average ratio. Since the sample CRDs' productions represent about 70% of total U.S. corn and soybean production as Table 2 shows, the 10-year moving average ratio is quite stable, around 0.97 for corn and 0.91 for soybeans.

Compared with methodology by Irwin et al. (2008), evaluating predictions of all CRDs is equivalent to evaluating a simple average of all forecasts without weight. The only difference lies on the weight. Without further knowledge of harvested acreage effect, the forecasting results of both pooled average CRD level predictions and the U.S. yield predictions will be reported.

6.3 Evaluation Standards

In order to measure forecasting accuracy, evaluation of forecast errors needs to be identified. Forecast error e_i in a specific region i was defined in a standard way, i.e. the difference between the actual yield y_i and the predicted yield $\hat{y}_i$. Several popular evaluation standards were adopted in this thesis. First of all, root mean squared error (RMSE) is a commonly used measure that has the same unit as the evaluating variables. It is defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2}$$

where *n* is the number of forecasts observations. Another measure similar to RMSE is root mean squared percentage error (RMSPE), which is often used to compare models with different units. Since RMSPE reports results in percentage term, it also provides the relative magnitude of forecast errors. Mathematically, RMSPE is calculated as

$$RMSPE = \sqrt{\frac{1}{n}\sum_i (100 \times \frac{y_i - \hat{y}_i}{y_i})^2}$$

Besides the above two measures, mean absolute error (MAE) is also adopted. It is simply the average of forecast errors without considering their direction. MAE is calculated as

$$MAE = \frac{1}{n}\sum_i \left| y_i - \hat{y}_i \right|$$

Likewise, mean absolute percentage error (MAPE) is the corresponding percentage measure to MAE. It is defined as

$$MAPE = \frac{1}{n}\sum_i \left| 100 \times \frac{y_i - \hat{y}_i}{y_i} \right|$$

6.4 Forecasting Results

In Chapter 4 and Chapter 5, three statistical models that utilize the entire data set were discussed in detail. Particularly, the modified Thompson and modified Roberts and Schlenker models specified a fixed effect term in the regression, which in essence assumed the same response coefficients for independent variables across all the CRDs. In acknowledgement of existence of heterogeneity among CRDs, another extreme case in which each CRD was assessed individually will be explored. Specifically, forecasting results were calculated for each individual CRD with the functional form of modified Thompson models and modified Roberts and Schlenker models.

Besides two CRD level models, previous studies by Irwin et al. (2009a, 2009b) had derived U.S. corn and soybean yield forecasts from state level predictions of Illinois, Indiana, and Iowa. To make a broad comparison, their method was also adopted in this thesis. Regression analysis was performed on these three states with the functional form of modified Thompson models. Apparently, since observations were on state level, there were no pooled average CRD forecasts from three-state model.

With forecasts from three models developed in Chapter 4 and Chapter 5, two CRD level models, and the three-state model, one would naturally think of deriving a composite forecast by averaging predictions of these six models. By taking the average, forecasting errors of opposite directions from different models might cancel out, yielding a more accurate prediction. Due to different data availability, composite forecasts were calculated for all 6 models for the period of 1988 through 2005, but calculated from 4 models except modified Roberts and Schlenker models and their CRD level models for the entire period up to 2010.

6.4.1 Corn Yield Forecasts

Forecast errors from all corn models are showed in Table 14 and plotted separately in Figure 32. While 1993 signified a major failure for all prediction models due to lasting unfavorable weather, the forecasting performances generally displayed a consistent pattern. Specifically, predictions tended to fluctuate up and down around the actual corn yield during the forecasting period, except for the modified Thompson corn model, in which predictions since 1998 had always been positively biased. Panel B of Figure 32 showed the forecast errors of modified Roberts and Schlenker corn model. Compared to other models, its errors were always within 8.5 bushels per acre except 1991 and 1993 until 2003. However, in 2004 and 2005, the errors jumped to 10.9 and 11.4 bushels per acre, whereas predictions from other models were below 10 bushels per acre. Owing to unavailability of data after 2005, it's hard to tell whether this strange behavior was due to a fundamental change in functional form or it was only poor predictions during these two years. Future works with a complete data set may help answer this question.

Table 15 presents out-of-sample forecasting accuracy statistics for corn models, and results from recursive method and resampling method were displayed in the top half and bottom half of the table, respectively. For the recursive method, the first round of competition was held for forecasts from 1988 through 2005. In forecasting area-weighted U.S. total corn yield, the composite prediction of all 6 models outperformed any single prediction in terms of all 4 criteria, meaning forecast errors from different models did compensate themselves. Besides the composite prediction, the best forecast came from panel GWR model in 3 criteria except for RMSPE, where the modified Roberts and Schlenker CRD level model beat all other models. If the forecasts were pooled together and averaged without weights, then the best prediction was from modified Roberts and Schlenker model by 3 measurements, and again their CRD level model won the competition in terms of RMSPE. Nonetheless, when the forecasting period was extended to 2010, the best U.S. corn yield prediction and the best average prediction were obtained from panel GWR model and modified Thompson model, respectively. By beating the composite forecast of 4 models, it indicated that forecast errors from other three models outweighed errors from panel GWR model.

The bottom half of Table 15 showed results from resampling method. In predicting area-weighted U.S. corn yield, the composite forecast stood out in most cases regardless of the size of the resampling pool. It was only beaten by modified Thompson model when bootstrapping from the entire period of 1960 through 2010 in terms of MAPE. The best average CRD forecast was from modified Thompson model, and it was beaten by modified Roberts and Schlenker model for 6-year resampling method in terms of RMSPE.

Compared with two benchmarks, none of these models outperformed USDA forecasts, but they were all better than trend-only forecasts. As Good and Irwin (2006) described, USDA utilized estimated planted and harvested acreage, a farmer-reported survey, and an objective yield survey to generate its crop production forecast. The procedure is highly sophisticated and takes advantage of both subjective and objective measurements. In contrast, regression models only absorb weather and planting progress data without farmer-provided information and real field data. Naturally, their predictive performance was worse than USDA forecast but better than trend-only model.

In sum, for area-weighted corn yield forecast, composite predictions won the most competitions followed by panel GWR model, and for average CRD forecast competition, modified Thompson model performed the best followed by modified Roberts and Schlenker model. An interesting finding is that in predicting area-weighted yield with recursive method, individual CRD level corn models tended to beat their corresponding fixed effect models in most cases, suggesting historical data of one specific district might be more informative than data from all CRDs in predicting yield of its own. But with randomly selected forecast years, fixed effect models yielded more accurate predictions. In addition, there is no denying that regression results from fixed models were valuable since they revealed the overall effect.

6.4.2 Soybean Yield Forecasts

Table 16 and Figure 33 showed out-of-sample forecast errors of U.S. soybean yield from different models. The standard deviations ranged from 2.2 to 2.9 bushels per acre, and most errors were within one standard deviation from the average. A couple of exceptions occurred in 1994 and 2003. While 1994 marked a high yield of 41.4 bushels per acre due to very favorable weather condition to soybeans production, 2003 happened to be the other way around with a record low yield of 33.9 bushels per acre since 1993.

Forecasting accuracy statistics for all soybean models were presented in Table 17. Similar to the table of corn models, the top half listed results from the recursive method. For the forecasting period of 1988 through 2005, the composite U.S. soybean yield prediction of 6 models was barely beaten by modified Thompson model in terms of MAE and beat all other models by the rest of criteria, and the best average CRD forecast was from modified Thompson model except for RMSPE criterion won by modified Roberts and Schlenker CRD level model. When including data of 2006 through 2010, the best area-weighted U.S. soybean yield prediction and average CRD prediction were consistently obtained from modified Thompson soybean model by all measurements.

The bottom half of Table 17 displayed competition results of resampling method. For area-weighted U.S. soybean yield forecast with 6-year resampling pool, the composite forecast appeared to be the best, followed by modified Thompson soybean model. Nonetheless, for 7-year resampling pool and both average CRD forecasts, modified Thompson soybean model won all the competitions. These results were similar to corn competitions. Again, the USDA forecasts turned out to be better than all regression predictions, and trend-only model underperformed most of 6 models.

Compared with results from corn models where the panel GWR model showed the strongest predictive power next to composite forecast, modified Thompson soybean model produced the most accurate forecasts for soybean yields besides composite forecast. Particularly, in predicting average CRD soybean yield, the modified Thompson fixed effect soybean model consistently yielded the best result regardless of time dimension. The more concise functional form of modified Thompson soybean model than that of corn model could possibly contribute to its better performance.

6.5 Summary

In comparing predictive power of different models, crop yield forecasting competitions were performed in this chapter. Two methods, a conventional recursive method and a resampling by year method, were applied to calculate harvest acreage-weighted U.S. crop yield and average CRD yield. Based on RMSE, RMSPE, MAE, and MAPE criteria, the composite forecast of 6 models and modified Thompson model appeared to produce the most accurate yields. Nonetheless, USDA forecasts were consistently better than predictions of all regression models and composite forecasts. Finally, since the growing degree days data were not available after 2005, further competition may be held for all model on the same time horizon.

7. Conclusion

7.1 Review

This thesis explored the relationship between corn and soybean yields and weather conditions. Previous work can be divided into two categories, i.e. agronomic study and empirical statistical study. Agronomists analyze plant physiology and phenology to draw conclusion about crop yields, whereas agricultural economists try to reveal the relationship from statistical evidence. As an empirical study, this thesis borrowed ideas from both sides and combined their advantages.

To build a valid model for corn and soybean yields in the U.S., the research subject was selected on crop reporting district level according to each CRD's production contribution. The final sampling pool included 49 original and one revised CRDs for corn models, and 38 original and 4 revised CRDs for soybean models.

Three statistical models were studied in detail. The modified Thompson model, as a fixed effect model, utilized the idea of vegetative and reproductive stages in plant physiology. Specifically, two different temperature variables captured the effect of heat during these two stages separately. In addition, the late planting variable was also found important in determining crop yield. Another fixed effect model, the modified Roberts and Schlenker model assumed cumulative and substitutable temperature effect over time, so growing degree days was used as a measurement of heat instead of monthly average temperatures. Nonetheless, due to an incomplete data set issue, the analysis might be limited. The last spatial model adopted panel geographically weighted regression technique and estimated coefficients for each single CRD.

The functional form was consistent with modified Thompson models so that estimations would include information up to 2010.

A forecasting performance competition was held to compare predictive power of all models. The evaluation was based on two methods, i.e. recursive method and resampling strategy. Since forecasts from modified Roberts and Schlenker models were not available after 2005, two sets of competitions with different sampling periods were performed respectively on U.S. total crop yield and average CRD yield.

7.2 Summary of Findings

Due to different underlying assumptions and specifications, three models served different purposes. Results from modified Thompson models showed that technology change contributed about 1.9 and 0.45 bushels per acre increase to corn and soybean yield every year, respectively. On the other hand, every 1% late planting of corn and soybeans would result in 0.25 and 0.05 bushels per acre decrease. While the sign of precipitation variables had indicated concave effects to crop yield, their impact during the reproductive stage was estimated to be greater compared to the vegetative stage. The same pattern was also found in pre-silking/flowering and post-silking/flowering temperature effects. As a result, one implication from modified Thompson models was that unusual weather conditions during July and August would affect crop yields to a greater magnitude than unfavorable weathers during early stages.

Modified Roberts and Schlenker models did not differentiate temperature effect during vegetative and reproductive stages, but identified moderate and extreme heat impact separately. Similar to results from modified Thompson models, the technology advancement was estimated to bring corn and soybean yield up by 1.7 and 0.42 bushels per acre every year. Precipitation was measured as one single variable during the entire growing season of March through August. Interestingly, the corn model was the only one that indicated the average total precipitation exceeded the optimal value that maximizes crop yield, whereas all other models suggested additional rainfall would be helpful. Finally, results showed that moderate heat was favorable to crop yield, but the impact was much smaller than the damage caused by extreme heat at the same magnitude.

The functional form of panel GWR models was adopted from modified Thompson models. Unsurprisingly, the average effects were generally consistent with outcomes from fixed effect models, although a few CRDs showed unexpected results. Another feature of panel GWR model was that coefficients were estimated exclusively for each CRD, so spatial patterns of yield response to weather variables can be studied. However, after examining plots of estimated coefficients and marginal effects of rainfall, it does not suggest any clear stable grouping patterns. Further effort to group homogeneous CRDs was not applicable.

With different models and specifications, the last part of this thesis provided a comprehensive forecasting competition to evaluate their predictive powers. Comparisons were made with 4 criteria, i.e. root mean squared error, root mean squared percentage error, mean absolute error, and mean absolute percentage error. The results indicated that in case of corn, the panel GWR corn model performed the best among all the models, and the modified Thompson soybean model and its CRD level model produced the most accurate forecasts for soybean yields.

7.3 Future Works

This study analyzed factors that affect crop yields, and mainly focused on technology trend, date of planting activities, and weather variables such as precipitation and temperature.

Yet, there is another very important aspect that undoubtedly has significant impact on crop yield – soil type. A recent study by Kravchenko and Bullock (2000) has found that soil properties explained about 5 to 71% of yield variations from field to field and from year to year. Nonetheless, due to high variability of soil types for even a small area and lack of uniform standard for soil properties, it is difficult to incorporate soil information to the regression analysis at this stage. Future studies may consider construction of an index for soil types.

Another direction for future research involves in panel data analysis. When working with time series data, there are several techniques that can identify structural change in one or more variables in a linear regression. For example, the Chow test is used to check pre-defined structural break, and Quandt likelihood ratio (QLR) test is able to recognize unknown break point. However, those tests are not applicable to panel data, in which the addition of cross sectional dimension complicates the statistical issue. Therefore, future studies may resort to improved econometric technique, and perform structural change test on selected variables for further hypothesis testing. One meaningful hypothesis would be if there has been an accelerated yield growth for corn and soybeans by testing structural change on trend variable.

Lastly, one limitation of this study was owing to incomplete growing degree days data set after 2005. This issue could be easily solved once the data set is updated. While the regression results may not change drastically, the ranking of forecasting performance may improve, since the modified Roberts and Schlenker models and its CRD level models predicted very well for the period of 1988 to 2005.

7.4 Concluding Remarks

This thesis established a set of benchmark crop-weather models that incorporate both agronomy knowledge and empirical statistical evidence. Especially, the treatment of temperature variables was the key point to each of these models. With different assumptions, results from three models explained crop, technology, and weather relations from different perspectives. Admittedly, the true relationship is still debatable.

The key findings of this study include: 1) the impact of heat and precipitation is greater during crops' reproductive stage than vegetative stage; 2) the beneficial effect of moderate heat to crop yield is much smaller than the damage caused by extreme heat at the same magnitude; and 3) crop-weather relations vary significantly across CRDs, and the heterogeneity issue prevents further aggregation to larger modeling regions. These conclusions are useful guidelines to yield forecasting activities. However, with ever-changing weather conditions, and improvement of technology and agricultural management, there is no doubt that the crop-weather study should continue being updated.

Table 1. The Percentage of Irrigated Corn in Nebraska CRDs, 1960 – 2010

Year	Northeast	Central	East	Southwest	South	Southeast
1960	5.94%	73.10%	29.23%	72.68%	83.07%	30.34%
1961	6.96%	75.02%	35.99%	71.30%	82.83%	33.87%
1962	5.24%	69.56%	31.16%	68.87%	76.83%	28.59%
1963	6.32%	80.49%	38.72%	69.81%	88.10%	31.29%
1964	7.54%	82.99%	47.01%	79.33%	93.72%	44.20%
1965	6.91%	85.12%	44.51%	82.70%	94.98%	44.58%
1966	7.23%	82.01%	45.77%	78.00%	94.25%	53.13%
1967	10.72%	87.79%	50.68%	84.83%	96.57%	60.84%
1968	20.04%	88.12%	54.87%	88.96%	96.42%	63.39%
1969	12.76%	87.02%	50.96%	86.18%	94.49%	58.35%
1970	18.94%	90.28%	62.04%	88.88%	95.29%	66.34%
1971	17.58%	91.17%	55.62%	88.03%	96.51%	61.81%
1972	13.89%	86.25%	52.70%	86.66%	95.70%	58.43%
1973	16.20%	83.05%	50.98%	85.96%	92.55%	53.62%
1974	36.55%	92.83%	77.01%	93.08%	96.52%	80.27%
1975	35.26%	94.36%	70.26%	94.68%	97.45%	71.97%
1976	47.55%	93.91%	79.38%	96.30%	97.95%	73.43%
1977	39.68%	91.39%	76.64%	95.39%	97.20%	85.13%
1978	41.77%	93.53%	70.15%	97.30%	98.69%	76.17%
1979	39.34%	93.00%	69.52%	96.74%	98.14%	73.54%
1980	51.10%	97.83%	74.36%	96.97%	98.96%	83.47%
1981	46.09%	96.40%	74.42%	97.08%	98.86%	77.97%
1982	37.46%	92.63%	64.12%	93.20%	96.15%	72.24%
1983	45.50%	93.48%	72.95%	94.69%	97.26%	84.54%
1984	42.40%	93.36%	72.19%	94.18%	98.25%	83.04%
1985	42.32%	92.19%	67.78%	94.57%	96.27%	75.61%
1986	44.73%	93.29%	66.45%	93.02%	96.56%	74.33%
1987	46.36%	93.54%	70.13%	94.10%	96.84%	78.31%
1988	54.82%	94.76%	73.45%	92.54%	96.36%	81.97%
1989	58.80%	95.69%	77.03%	93.15%	97.41%	86.98%
1990	47.28%	94.74%	70.28%	94.61%	97.25%	75.32%
1991	52.34%	95.58%	71.47%	92.74%	96.54%	77.55%
1992	41.43%	91.18%	63.54%	90.52%	94.06%	71.68%
1993	41.38%	90.26%	60.62%	88.30%	89.60%	70.21%
1994	43.02%	91.50%	62.50%	89.79%	91.08%	63.36%
1995	48.71%	93.76%	71.04%	92.65%	93.38%	71.75%
1996	46.60%	90.58%	66.00%	88.78%	90.72%	64.96%
1997	43.13%	90.97%	62.07%	87.79%	89.29%	63.40%
1998	41.02%	88.74%	59.09%	84.74%	84.83%	56.54%
1999	38.20%	87.96%	56.51%	81.75%	82.64%	55.51%
2000	47.12%	96.15%	61.58%	92.18%	88.28%	57.86%
2001	41.04%	92.52%	61.41%	86.48%	86.11%	52.78%
2002	52.26%	99.20%	75.13%	97.61%	98.42%	75.96%
2003	48.74%	97.20%	72.08%	95.44%	95.21%	68.32%
2004	42.36%	95.26%	59.18%	89.81%	90.19%	50.31%
2005	45.40%	94.94%	65.09%	88.41%	87.07%	55.18%
2006	51.56%	95.54%	64.88%	91.24%	89.40%	56.00%
2007	48.63%	91.84%	63.00%	82.53%	85.06%	55.13%
2008	48.00%	91.05%	61.38%	81.18%	81.36%	52.13%
2009	45.29%	89.41%	60.02%	78.69%	81.96%	50.29%
2010	43.72%	89.03%	59.90%	76.49%	81.46%	51.11%
Mean	35.75%	90.42%	61.82%	88.25%	92.43%	63.59%

Table 2. The Percentage of Sample CRDs' Crop Production to the U.S. Total Production, 1960 – 2010

Year	Corn	Soybeans
1960	73.93%	71.23%
1961	74.58%	74.21%
1962	76.74%	72.93%
1963	77.14%	74.21%
1964	76.81%	71.13%
1965	76.81%	69.64%
1966	78.76%	67.59%
1967	76.49%	63.81%
1968	76.76%	68.47%
1969	77.47%	67.72%
1970	75.91%	67.79%
1971	76.60%	67.39%
1972	76.71%	69.96%
1973	75.12%	68.72%
1974	70.15%	65.05%
1975	73.59%	65.62%
1976	69.97%	64.94%
1977	72.47%	67.00%
1978	73.35%	64.99%
1979	73.38%	61.65%
1980	74.47%	69.94%
1981	72.84%	63.93%
1982	73.16%	62.39%
1983	67.68%	64.92%
1984	70.68%	62.47%
1985	72.37%	67.89%
1986	73.02%	71.40%
1987	72.03%	70.54%
1988	66.88%	63.02%
1989	72.83%	68.18%
1990	73.67%	69.45%
1991	70.88%	68.89%
1992	73.57%	67.96%
1993	70.01%	69.39%
1994	73.07%	68.15%
1995	70.43%	70.87%
1996	69.02%	67.09%
1997	69.56%	67.25%
1998	71.09%	68.73%
1999	71.23%	67.01%
2000	70.49%	67.11%
2001	71.16%	65.95%
2002	72.94%	67.15%
2003	72.24%	61.40%
2004	72.65%	64.05%
2005	71.61%	63.89%
2006	73.58%	64.20%
2007	69.37%	63.12%
2008	69.44%	60.08%
2009	69.85%	58.97%
2010	68.77%	60.44%
Mean	72.81%	66.86%

Table 3. Summary Statistics for Major Corn Production CRDs, 1960 – 2005/2010

Variable	Mean	Median	Minimum	Maximum	Range	Standard Deviation
1960 – 2010 data						
Yield (bushels per acre)	109.55	108.6	26.5	195	168.5	33.88
Late planting (percent)	18.93	14.29	0	85.86	85.86	16.92
Precipitation (inches)						
Pre-season	16.2	15.97	3.51	44.51	41	5.8
April	3.45	3.26	0.14	11.27	11.13	1.59
May	4.09	3.88	0.45	11.94	11.49	1.88
June	4.23	3.94	0.36	13.28	12.92	1.92
July	3.99	3.72	0.32	16	15.68	1.85
August	3.68	3.38	0.37	16.66	16.29	1.89
Pre-silking	11.77	11.5	3.36	25.22	21.86	3.49
Post-silking	7.67	7.35	1.72	23.58	21.86	2.63
Temperature (degrees Fahrenheit)						
April	49.61	49.6	35.5	62.8	27.3	4.55
May	60.27	60	49.6	73.3	23.7	4.2
June	69.67	69.6	59.2	80.2	21	3.13
July	73.67	73.6	63.3	83.5	20.2	2.99
August	71.58	71.4	62.4	83.8	21.4	3.28
Pre-silking	59.85	59.67	51.33	70.8	19.47	3.26
Post-silking	72.63	72.4	63.6	82	18.4	2.83
1960 – 2005 data						
Total precipitation (inches)	54.81	53.92	19.31	111.51	92.2	11.56
Growing degree days						
Moderate (10 to 29˚C)	1400.45	1386.54	952.43	1996.91	1044.48	184.95
Extreme (above 29˚C)	26.48	21.08	1.09	126.43	125.34	20.28

Table 4. Summary Statistics for Major Soybean Production CRDs, 1960 – 2005/2010

Variable	Mean	Median	Minimum	Maximum	Range	Standard Deviation
1960 – 2010 data						
Yield (bushels per acre)	35.43	35	10.9	57	46.1	8.62
Late planting (percent)	30.13	24	0	93.71	93.71	22.79
Precipitation (inches)						
Pre-season	20.95	20.6	6.6	53.49	46.89	6.72
May	4.26	4.05	0.45	13.72	13.27	1.94
June	4.28	3.98	0.36	13.28	12.92	1.92
July	4.08	3.81	0.35	16	15.65	1.9
August	3.69	3.38	0.33	13.15	12.82	1.88
Pre-flowering	8.54	8.28	1.32	20.82	19.5	2.92
Post-flowering	7.77	7.44	1.76	23.58	21.82	2.66
Temperature (degrees Fahrenheit)						
May	61.09	60.7	50.6	75.5	24.9	4.33
June	70.38	70.3	59.2	82.4	23.2	3.17
July	74.24	74.1	63.6	85.8	22.2	3.08
August	72.19	72	62.4	85.2	22.8	3.46
Pre-flowering	65.74	65.35	57.55	76.35	18.8	3.19
Post-flowering	73.21	72.95	63.8	84.05	20.25	2.99
1960 – 2005 data						
Total precipitation (cm)	56.42	55.72	18.93	111.15	92.22	11.86
Growing degree days						
Moderate (10 to 30˚C)	1466.39	1440.4	957.1	2184.2	1227.1	209.95
Extreme (above 30˚C)	18.41	13.42	0.22	143.58	143.36	16.86

Table 5. Data Structure of Panel GWR for Coefficient Estimate of Region A

Year	Y	X	Weight
1	Y_{A1}	X_{A1}	1
2	Y_{A2}	X_{A2}	1
3	Y_{A3}	X_{A3}	1
1	Y_{B1}	X_{B1}	w_{AB}
2	Y_{B2}	X_{B2}	w_{AB}
3	Y_{B3}	X_{B3}	w_{AB}
1	Y_{C1}	X_{C1}	w_{AC}
2	Y_{C2}	X_{C2}	w_{AC}
3	Y_{C3}	X_{C3}	w_{AC}

Table 6. Data Structure of Panel GWR Model

Year	Y	X_A	X_B	X_C	Weight
1	Y_{A1}	X_{A1}	0	0	1
2	Y_{A2}	X_{A2}	0	0	1
3	Y_{A3}	X_{A3}	0	0	1
1	Y_{B1}	X_{B1}	0	0	w_{AB}
2	Y_{B2}	X_{B2}	0	0	w_{AB}
3	Y_{B3}	X_{B3}	0	0	w_{AB}
1	Y_{C1}	X_{C1}	0	0	w_{AC}
2	Y_{C2}	X_{C2}	0	0	w_{AC}
3	Y_{C3}	X_{C3}	0	0	w_{AC}
1	Y_{A1}	0	X_{A1}	0	w_{AB}
2	Y_{A2}	0	X_{A2}	0	w_{AB}
3	Y_{A3}	0	X_{A3}	0	w_{AB}
1	Y_{B1}	0	X_{B1}	0	1
2	Y_{B2}	0	X_{B2}	0	1
3	Y_{B3}	0	X_{B3}	0	1
1	Y_{C1}	0	X_{C1}	0	w_{BC}
2	Y_{C2}	0	X_{C2}	0	w_{BC}
3	Y_{C3}	0	X_{C3}	0	w_{BC}
1	Y_{A1}	0	0	X_{A1}	w_{AC}
2	Y_{A2}	0	0	X_{A2}	w_{AC}
3	Y_{A3}	0	0	X_{A3}	w_{AC}
1	Y_{B1}	0	0	X_{B1}	w_{BC}
2	Y_{B2}	0	0	X_{B2}	w_{BC}
3	Y_{B3}	0	0	X_{B3}	w_{BC}
1	Y_{C1}	0	0	X_{C1}	1
2	Y_{C2}	0	0	X_{C2}	1
3	Y_{C3}	0	0	X_{C3}	1

Table 7. Coefficient Estimates of Modified Thompson Model for Corn, 1960 – 2010

Independent variable	Coefficient (Robust Std. Err.)	t-stat (p-value)
Constant	235.887*** (33.140)	7.12 (0.000)
Trend	1.864*** (0.054)	34.32 (0.000)
Late planting	-0.254*** (0.045)	-5.65 (0.000)
Pre-season precipitation	1.224*** (0.380)	3.23 (0.009)
Pre-season precipitation^2	-0.031** (0.010)	-3.07 (0.012)
April precipitation	2.449*** (0.617)	3.96 (0.003)
April precipitation^2	-0.227*** (0.064)	-3.57 (0.005)
May precipitation	2.465*** (0.646)	3.82 (0.003)
May precipitation^2	-0.267*** (0.057)	-4.68 (0.001)
June precipitation	6.373*** (0.835)	7.63 (0.000)
June precipitation^2	-0.591*** (0.097)	-6.09 (0.000)
July precipitation	9.853*** (0.417)	23.62 (0.000)
July precipitation^2	-0.740*** (0.037)	-19.90 (0.000)
August precipitation	2.104*** (0.450)	4.67 (0.001)
August precipitation^2	-0.158*** (0.040)	-3.92 (0.003)
Pre-silking temperature	0.206 (0.130)	1.59 (0.144)
Post-silking temperature	-3.407*** (0.428)	-7.97 (0.000)
R^2	0.884	
Adjusted R^2	0.881	
Standard error	11.69	
F-stat (degree of freedom)	839.02*** (19, 2481)	

Note: one, two, and three asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8. Coefficient Estimates of Modified Thompson Model for Soybeans, 1960 – 2010

Independent variable	Coefficient (Robust Std. Err.)	t-stat (p-value)
Constant	15.761 (12.101)	1.30 (0.2290)
Trend	0.450*** (0.017)	27.10 (0.000)
Late planting	-0.048*** (0.006)	-7.69 (0.000)
Pre-season precipitation	0.110* (0.049)	2.25 (0.055)
Pre-season precipitation^2	-0.002* (0.001)	-2.08 (0.071)
Pre-flowering precipitation	1.470*** (0.279)	5.26 (0.001)
Pre-flowering precipitation^2	-0.074*** (0.016)	-4.75 (0.001)
Post-flowering precipitation	2.741*** (0.291)	9.41 (0.000)
Post-flowering precipitation^2	-0.121*** (0.016)	-7.49 (0.000)
Pre- flowering temperature	0.240*** (0.044)	5.52 (0.001)
Post- flowering temperature	-0.376** (0.136)	-2.76 (0.025)
R^2	0.838	
Adjusted R^2	0.834	
Standard error	3.51	
F-stat (degree of freedom)	855.49*** (10, 2090)	

Note: one, two, and three asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 9. Coefficient Estimates of Modified Robert and Schlenker Model for Corn and Soybeans, 1960 – 2005

Independent variable	Corn		Soybeans	
	Coefficient (Robust Std. Err.)	t-stat (p-value)	Coefficient (Robust Std. Err.)	t-stat (p-value)
Constant	-10.78	-0.67	-5.575	-0.95
	(16.10)	(0.518)	(5.878)	(0.371)
Trend	1.71***	30.92	0.422***	22.90
	(0.06)	(0.000)	(0.018)	(0.000)
Total precipitation	2.03***	5.93	0.517***	5.62
	(0.34)	(0.000)	(0.092)	(0.001)
Total precipitation^2	-0.02***	-6.42	-0.004***	-5.66
	(0.003)	(0.000)	(0.001)	(0.000)
Moderate temperature	0.03***	3.80	0.013***	4.75
	(0.01)	(0.003)	(0.003)	(0.001)
Extreme temperature	-0.75***	-24.91	-0.184***	-12.84
	(0.03)	(0.000)	(0.014)	(0.000)
R^2	0.845		0.779	
Adjusted R^2	0.841		0.774	
Standard error	12.31		3.81	
F-stat	1963.69***		995.11***	
(degree of freedom)	(5, 2245)		(5, 1885)	

Note: one, two, and three asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 10. Coefficient Estimates of Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

CRD	Trend	Late Planting	Pre-season precipitation	Pre-season precipitation squared	April precipitation	April precipitation squared	May precipitation	May precipitation squared	June precipitation	June precipitation squared	July precipitation	July precipitation squared	August precipitation	August precipitation squared	Pre-silking temperature	Post-silking temperature
IL-10	1.960***	-0.123*	1.528	-0.036	5.472*	-0.622*	0.412	-0.142	11.110***	-1.106***	10.358***	-0.746***	-1.031	0.081	1.882***	-4.142***
IL-20	1.728***	-0.295***	-0.935	0.041	0.672	0.028	5.208***	-0.641***	11.144***	-1.083***	8.944***	-0.534**	1.085	-0.012	1.468***	-3.642***
IL-30	1.971***	-0.314***	5.590***	-0.156***	3.275	-0.272	0.350	-0.066	11.009***	-1.089***	15.026***	-1.157***	5.335***	-0.608***	1.143***	-4.253***
IL-40	1.911***	-0.168**	2.962*	-0.077*	6.215***	-0.638**	-1.448	0.027	11.912***	-1.235***	10.793***	-0.741***	2.060	-0.310	1.691***	-4.156***
IL-50	1.751***	-0.166**	3.320**	-0.086**	9.250***	-1.001***	0.553	-0.184	10.343***	-1.009***	11.633***	-0.772***	2.071	-0.175	1.633***	-4.367***
IL-60	2.019***	-0.186**	3.029**	-0.084**	3.714*	-0.297	-0.290	-0.068	10.203***	-1.068***	11.260***	-0.822***	3.614*	-0.511**	1.494***	-3.977***
IL-70	1.807***	-0.040	2.145*	-0.051**	3.042	-0.319**	0.782	-0.181	6.831***	-0.778***	12.511***	-0.770***	3.399	-0.128	0.413	-3.185***
IL-80	1.750***	-0.066	1.185	-0.021	1.345	-0.102	1.700	-0.222*	8.035***	-0.873***	12.013***	-0.772**	5.185**	-0.305	0.639	-3.478***
IL-90	1.780***	-0.045	1.188	-0.025	2.033	-0.192	1.649	-0.241*	8.200***	-0.903***	12.262***	-0.757***	4.161*	-0.185	0.336	-3.163***
IN-10	1.698***	-0.134**	3.177*	-0.071	4.426	-0.477	2.897	-0.359	8.592***	-0.733***	10.347***	-0.636***	3.094	-0.224	1.015**	-3.946***
IN-20	1.735***	-0.133**	2.093	-0.042	2.817	-0.306	4.819	-0.594	7.770***	-0.650***	9.382***	-0.566***	3.066	-0.202	0.654	-3.520***
IN-30	1.743***	-0.168***	1.650	-0.041	2.654	-0.256	5.825**	-0.707**	6.807***	-0.603***	9.193***	-0.592***	3.748*	-0.262	0.551	-3.338***
IN-40	1.695***	-0.120*	2.350	-0.052	8.555***	-0.929***	0.561	-0.123	9.419***	-0.803***	12.251***	-0.824***	2.183	-0.153	1.492***	-4.250***
IN-50	1.683***	-0.162**	3.215*	-0.072*	6.921***	-0.722***	1.186	-0.130	7.763***	-0.608***	8.781***	-0.569***	6.632**	-0.651**	1.441***	-4.307***
IN-60	1.682***	-0.204***	4.086***	-0.098**	5.576***	-0.565***	2.186	-0.218	6.301***	-0.469**	6.776***	-0.430***	9.098***	-0.920***	1.347***	-4.304***
IN-70	1.848***	-0.034	-0.131	-0.001	1.006	-0.138	0.728	-0.159	9.297***	-1.046***	13.469***	-0.890***	2.636	-0.118	0.241	-2.698***
IA-10	2.137***	-0.589***	0.766	0.017	3.293	-0.322	3.002	-0.255	3.745**	-0.388***	10.002***	-0.955***	1.381	-0.099	0.354	-2.813***
IA-20	1.974***	-0.597***	3.936**	-0.133**	2.526	-0.184	2.376	-0.208	3.952**	-0.385**	8.938***	-0.838***	0.760	-0.079	-0.026	-2.525***
IA-30	2.000***	-0.528***	2.128	-0.065	1.972	-0.115	4.543**	-0.400**	9.470***	-0.812***	8.155***	-0.633***	0.767	-0.070	0.638	-3.261***
IA-40	2.015***	-0.419***	2.281**	-0.052	3.938*	-0.387	0.246	-0.002	1.449	-0.114	9.044***	-0.803***	2.992**	-0.309**	0.426	-2.852***
IA-50	1.997***	-0.540***	4.894***	-0.170***	4.268*	-0.450	5.101**	-0.517**	7.869***	-0.715***	8.439***	-0.669***	4.730***	-0.420***	0.128	-3.152***
IA-60	1.944***	-0.422***	4.169***	-0.129**	2.024	-0.190	2.948*	-0.285*	10.216***	-0.917***	10.795***	-0.811***	1.374	-0.148	0.858**	-3.673***
IA-70	1.910***	-0.347***	2.276**	-0.049	5.650**	-0.558**	-1.019	0.138	1.736	-0.108	9.174***	-0.783***	3.005**	-0.312***	0.866**	-3.257***
IA-80	1.892***	-0.336***	3.079***	-0.108***	4.472*	-0.504	4.247**	-0.467***	9.756***	-0.913***	10.480***	-0.745***	5.920***	-0.510***	0.707*	-3.679***
IA-90	1.881***	-0.342***	6.442***	-0.198***	3.122	-0.419	1.475	-0.171	11.265***	-1.062***	14.607***	-1.069***	2.933*	-0.261*	0.689*	-3.923***
KY-20	1.889***	-0.006	-0.601	0.007	0.560	-0.100	0.630	-0.166	10.187***	-1.115***	12.734***	-0.833***	2.910	-0.178	0.159	-2.565***
MI-70	1.689***	-0.256***	-5.190***	0.142***	-8.982**	1.236**	5.881***	-0.610**	5.998**	-0.479	13.006***	-1.127***	4.081**	-0.287	-0.852*	-1.291***
MI-80	1.694***	-0.264***	-4.782***	0.129***	-8.634**	1.220**	5.893***	-0.613**	6.829***	-0.595**	11.470***	-0.944**	3.976**	-0.294	-0.809*	-1.338***
MI-90	1.705***	-0.256***	-5.349***	0.150**	-4.987	0.578	2.240	-0.284	14.246***	-1.551***	9.664**	-0.860	4.547**	-0.440	-1.790***	-0.531
MN-40	2.035***	-0.417***	-0.612	0.083	3.130	-0.163	13.259***	-1.664***	6.572**	-0.760***	10.241***	-0.916***	-2.946	0.569*	-0.801**	-2.253***
MN-50	2.035***	-0.417***	-0.612	0.083	3.130	-0.163	13.259***	-1.664***	6.572**	-0.760***	10.241***	-0.916***	-2.946	0.569*	-0.801**	-2.253***
MN-70	2.168***	-0.563***	-0.473	0.081	4.169*	-0.489	4.767*	-0.454	10.348***	-1.123***	7.439***	-0.607***	-1.569	0.239	0.358	-2.888***
MN-80	1.992***	-0.471***	2.439	-0.093	-0.063	0.097	2.527	-0.309	6.014***	-0.645***	10.866***	-1.084***	-1.689	0.115	0.081	-2.408***
MN-90	1.978***	-0.405***	-1.082	0.029	-4.911*	0.555	3.902	-0.390	7.140***	-0.636***	6.784***	-0.447**	3.301**	-0.266**	-0.361	-1.984***
MO-12	1.756***	-0.229***	2.341**	-0.073**	5.020**	-0.608*	1.376	-0.194	9.324***	-0.937***	10.173***	-0.701***	2.353*	-0.207*	1.384***	-4.011***
MO-30	1.769***	-0.402***	5.925***	-0.173***	3.337	-0.295	0.065	-0.002	11.711***	-1.096***	17.716***	-1.270***	4.150**	-0.335*	0.665	-4.056***

Note: one, two, and three asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 10. Coefficient Estimates of Panel Geographically Weighted Regression Model for Corn, 1960 – 2010 (Continued)

CRD	Trend	Late Planting	Pre-season precipitation	Pre-season precipitation squared	April precipitation	April precipitation squared	May precipitation	May precipitation squared	June precipitation	June precipitation squared	July precipitation	July precipitation squared	August precipitation	August precipitation squared	Pre-silking temperature	Post-silking temperature
NE-30	1.802***	-0.187**	3.025**	-0.051	5.425***	-0.530*	-3.040	0.475**	1.010	0.099	17.725***	-1.784***	4.840**	-0.446	0.221	-3.332***
NE-60	1.552***	-0.249**	5.060***	-0.135**	5.062**	-0.401	-2.315	0.358*	1.478	0.035	12.073***	-0.815***	3.088	-0.316	0.722*	-3.685***
NE-90	1.512***	-0.246**	5.356***	-0.150**	5.353**	-0.434	-1.146	0.244	2.153	-0.030	12.126***	-0.791***	3.625*	-0.392*	0.837*	-3.863***
OH-10	1.852***	-0.219***	4.128**	-0.123**	4.595*	-0.553	5.412*	-0.692**	6.177***	-0.546**	12.043***	-0.955***	6.816***	-0.596***	0.547	-3.696***
OH-20	1.823***	-0.144**	1.075	-0.040	0.478	-0.181	4.191*	-0.558**	7.691***	-0.785***	11.135***	-0.809***	5.181***	-0.361*	0.996**	-3.550***
OH-40	1.824***	-0.218***	4.322***	-0.114***	4.115*	-0.511**	2.218	-0.247	6.137**	-0.494*	9.297***	-0.692***	9.520***	-0.949***	1.282***	-4.279***
OH-50	1.829***	-0.148**	2.129	-0.067	1.742	-0.368	4.455*	-0.576**	7.070***	-0.682**	10.041***	-0.719***	6.599***	-0.525**	1.290***	-3.959***
OH-70	1.827***	-0.202***	4.908***	-0.122***	3.482	-0.444*	2.128	-0.200	5.578**	-0.425	7.882***	-0.546***	10.397***	-1.148***	1.475***	-4.486***
SD-60	2.147***	-0.430***	1.151	0.003	-0.450	0.395	2.643	-0.012	10.054***	-1.101***	8.538***	-0.736**	0.464	0.003	-0.706*	-2.432***
SD-90	2.019***	-0.320***	2.218	-0.056	2.249	-0.093	0.603	0.165	3.938**	-0.299	12.974***	-1.324***	1.953	-0.092	-0.313	-2.774***
WI-40	1.975***	-0.369***	-1.785	0.053	-4.803	0.501	3.738	-0.391	7.197***	-0.641***	6.582***	-0.420**	3.476**	-0.282**	-0.428	-1.869***
WI-70	1.716***	-0.368***	1.046	-0.032	-4.078	0.495*	5.581***	-0.468***	10.760***	-0.887***	6.709***	-0.546***	0.637	-0.005	0.218	-2.641***
WI-80	1.547***	-0.302***	-2.085	0.071	-5.068*	0.530	6.086***	-0.557***	7.038***	-0.540***	3.540	-0.146	1.883	-0.020	0.393	-2.290***
WI-90	1.567***	-0.360***	-2.977**	0.103**	-4.183	0.441	6.427***	-0.611***	7.343***	-0.584***	3.768	-0.130	1.617	0.006	0.362	-2.206***
Mean	1.844	-0.279	1.720	-0.041	2.079	-0.184	2.816	-0.312	7.655	-0.721	10.388	-0.780	3.129	-0.251	0.524	-3.210

Note: one, two, and three asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 11. Marginal Effect at the Mean Value of Precipitation Variables in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

CRD	Pre-season Precipitation	April Precipitation	May Precipitation	June Precipitation	July Precipitation	August Precipitation
IL-10	0.352	0.849	-0.753	1.314	4.379	-0.369
IL-20	0.459	0.877	0.283	2.197	4.735	0.991
IL-30	0.050	1.177	-0.243	1.690	5.211	0.696
IL-40	0.173	1.477	-1.224	2.132	4.863	-0.122
IL-50	0.194	1.786	-0.979	2.109	4.975	0.793
IL-60	-0.206	1.439	-0.887	1.506	5.027	0.219
IL-70	-0.012	0.526	-0.862	0.372	6.122	2.567
IL-80	0.193	0.501	-0.409	1.091	6.007	3.150
IL-90	-0.075	0.327	-0.747	0.893	6.277	2.965
IN-10	0.482	0.884	0.156	2.358	5.050	1.394
IN-20	0.504	0.630	0.176	2.469	4.629	1.513
IN-30	0.156	0.836	0.225	2.003	4.646	1.846
IN-40	0.250	1.225	-0.550	2.376	4.828	1.035
IN-50	0.259	1.224	0.026	2.682	3.681	1.949
IN-60	0.327	1.258	0.316	2.340	3.165	2.642
IN-70	-0.198	-0.193	-0.934	0.780	5.598	1.813
IA-10	1.139	1.424	1.152	0.350	2.813	0.632
IA-20	0.690	1.301	0.604	0.204	1.437	0.109
IA-30	0.353	1.152	1.173	1.646	2.412	0.156
IA-40	1.047	1.493	0.231	0.392	2.756	0.602
IA-50	0.457	1.090	0.493	0.656	2.598	1.097
IA-60	0.296	0.693	0.571	1.518	3.730	0.097
IA-70	0.968	1.896	0.298	0.684	2.237	0.527
IA-80	-0.112	0.776	-0.099	1.031	3.764	1.576
IA-90	-0.001	0.026	-0.071	1.630	5.045	0.759
KY-20	-0.215	-0.380	-1.014	1.176	5.794	1.741
MI-70	0.408	-0.612	1.674	2.544	5.254	2.026
MI-80	-0.470	-1.136	1.716	2.531	5.292	1.958
MI-90	-0.416	-1.528	0.467	3.397	4.227	1.708
MN-40	0.946	2.383	3.573	0.607	3.892	0.542
MN-50	1.184	2.298	2.214	-0.095	3.559	1.405
MN-70	1.185	1.462	1.742	1.214	3.058	-0.010
MN-80	0.277	0.510	0.202	0.326	1.848	-0.758
MN-90	-0.343	-1.487	0.866	1.512	3.037	0.990
MO-12	-0.108	0.534	-0.539	0.207	3.986	0.693
MO-30	-0.778	1.043	0.047	2.557	7.220	1.775
NE-30	2.026	2.531	0.679	1.860	6.248	1.979
NE-60	2.093	2.802	0.819	1.787	6.517	0.890
NE-90	1.797	2.902	1.032	1.897	5.836	0.837
OH-10	-0.117	0.947	0.273	2.166	4.931	2.891
OH-20	-0.369	-0.711	0.033	1.646	4.951	2.682
OH-40	0.078	0.449	0.189	2.237	3.452	2.915
OH-50	-0.404	-0.874	-0.382	1.618	4.025	2.824
OH-70	-0.302	-0.004	0.267	2.190	3.357	2.425
SD-60	1.193	1.388	2.570	1.620	4.064	0.478
SD-90	1.221	1.751	1.739	1.634	4.751	1.427
WI-40	-0.396	-1.778	0.711	1.528	3.130	1.019
WI-70	0.164	-0.641	1.971	2.742	2.127	0.592
WI-80	0.022	-1.482	2.120	2.323	2.383	1.719
WI-90	0.199	-1.206	2.381	2.778	2.799	1.665
Mean	0.332	0.677	0.466	1.608	4.234	1.301

Table 12. Coefficient Estimates of Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

CRD	Trend	Late Planting	Pre-season precipitation	Pre-season precipitation^2	Pre-flowering precipitation	Pre-flowering precipitation^2	Post-flowering precipitation	Post-flowering precipitation^2	Pre-flowering temperature	Post-flowering temperature
IL-10	0.433***	-0.026*	-0.012	0.001	1.194***	-0.057***	2.304***	-0.102***	0.467***	-0.407***
IL-20	0.418***	-0.037***	-0.161	0.004	1.209***	-0.058**	3.047***	-0.139***	0.429***	-0.409***
IL-30	0.450***	-0.056***	0.094	-0.003	0.840***	-0.035**	2.241***	-0.097***	0.455***	-0.402***
IL-40	0.439***	-0.044***	-0.003	-0.002	0.730**	-0.030	3.616***	-0.170***	0.322***	-0.315***
IL-50	0.430***	-0.046***	0.029	-0.001	1.222***	-0.061***	4.045***	-0.193***	0.309***	-0.380***
IL-60	0.435***	-0.087***	0.074	-0.003	0.698**	-0.019	3.410***	-0.156***	0.354***	-0.373***
IL-70	0.439***	-0.029**	0.527***	-0.012***	1.549***	-0.084***	3.691***	-0.164***	0.165	-0.372***
IL-80	0.374***	-0.014	0.518***	-0.012***	1.073***	-0.057***	3.623***	-0.171***	0.193*	-0.388***
IL-90	0.398***	-0.020	0.353**	-0.008**	1.250***	-0.063***	3.406***	-0.144***	0.210*	-0.398***
IN-10	0.454***	-0.060***	-0.264	0.005	2.089***	-0.104***	4.238***	-0.206***	0.190*	-0.312***
IN-20	0.475***	-0.074***	-0.047	0.000	2.652***	-0.130***	4.283***	-0.210***	0.054	-0.263***
IN-30	0.448***	-0.069***	0.058	-0.001	2.989***	-0.159***	4.144***	-0.198***	0.107	-0.344***
IN-40	0.464***	-0.063***	-0.055	0.002	2.201***	-0.108***	3.577***	-0.165***	0.195*	-0.326***
IN-50	0.470***	-0.083***	0.124	-0.002	2.646***	-0.123***	3.326***	-0.157***	0.238**	-0.395***
IN-60	0.460***	-0.077***	0.216	-0.004	2.968***	-0.148***	3.403***	-0.160***	0.316***	-0.515***
IN-70	0.433***	-0.029**	0.341*	-0.007*	1.664***	-0.083***	3.308***	-0.138***	0.201*	-0.393***
IA-10	0.506***	-0.051***	-0.127	0.001	2.407***	-0.141***	2.319***	-0.117***	0.222**	-0.255***
IA-20	0.498***	-0.055***	0.236	-0.002	1.047**	-0.061***	1.967***	-0.086***	0.175*	-0.208**
IA-30	0.531***	-0.011	-0.108	0.005	0.903***	-0.042**	2.441***	-0.105***	0.442***	-0.455***
IA-40	0.476***	-0.046***	0.073	-0.001	1.533***	-0.081***	2.201***	-0.102***	0.352***	-0.358***
IA-50	0.478***	-0.043***	0.461***	-0.005	1.729***	-0.091***	2.348***	-0.097***	0.279**	-0.417***
IA-60	0.472***	-0.008	0.304**	-0.004*	1.235***	-0.061***	2.098***	-0.091***	0.494***	-0.509***
IA-70	0.455***	-0.065***	0.543***	-0.009***	1.648***	-0.077***	2.107***	-0.091***	0.422***	-0.523***
IA-80	0.445***	-0.100***	0.306**	-0.003	1.633***	-0.080***	2.269***	-0.090***	0.511***	-0.596***
IA-90	0.454***	-0.025*	0.369***	-0.006***	1.004***	-0.051***	1.768***	-0.075***	0.568***	-0.539***
MI-90	0.373***	-0.062***	-0.091	0.003	2.289***	-0.137***	2.753***	-0.109***	0.079	-0.183*
MN-40	0.514***	-0.028	0.046	0.000	3.168***	-0.208***	3.088***	-0.150***	-0.171*	-0.134
MN-50	0.514***	-0.052***	-0.087	0.001	2.727***	-0.168***	3.270***	-0.157***	-0.030	-0.196**
MN-70	0.509***	-0.040**	-0.371***	0.005*	3.155***	-0.197***	2.166***	-0.111***	0.148	-0.191**
MN-80	0.513***	-0.062***	0.040	-0.003	1.673***	-0.106***	2.499***	-0.124***	0.103	-0.158*
MN-90	0.573***	-0.038**	0.229	-0.005	1.148***	-0.070***	3.018***	-0.131***	0.122	-0.280***
MO-12	0.389***	-0.091***	0.081	-0.002	1.128***	-0.054***	2.476***	-0.102***	0.614***	-0.618***
MO-36	0.429***	-0.137***	0.009	-0.001	0.816**	-0.013	3.338***	-0.153***	0.472***	-0.488***
MO-45	0.324***	-0.039***	0.072	-0.002	0.630**	-0.025*	2.752***	-0.113***	0.721***	-0.719***
MO-90	0.391***	-0.019	0.044	-0.001	-0.083	0.010	4.469***	-0.260***	0.371***	-0.461***
NE-60	0.487***	-0.079***	0.155	-0.003	1.925***	-0.086***	2.773***	-0.118***	0.229**	-0.355***
OH-10	0.434***	-0.066***	0.171	-0.002	2.965***	-0.172***	3.712***	-0.173***	0.174	-0.399***
OH-20	0.445***	-0.069***	-0.152	0.002	3.264***	-0.191***	2.885***	-0.125***	0.289***	-0.420***
OH-40	0.457***	-0.081***	0.229	-0.005	3.123***	-0.164***	4.204***	-0.204***	0.340***	-0.595***
OH-50	0.468***	-0.083***	0.187	-0.004	3.212***	-0.171***	3.655***	-0.170***	0.380***	-0.612***
OH-70	0.463***	-0.061***	0.165	-0.004	2.860***	-0.148***	4.352***	-0.213***	0.409***	-0.644***
TN-12	0.392***	-0.020	-0.045	0.001	-0.928**	0.063***	4.647***	-0.275***	0.508***	-0.553***
Mean	0.453	-0.053	0.108	-0.002	1.743	-0.092	3.125	-0.146	0.296	-0.401

Note: one, two, and three asterisks denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 13. Marginal Effect at the Mean Value of Precipitation Variables in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

CRD	Pre-season precipitation	Pre-flowering precipitation	Post-flowering precipitation
IL-10	0.034	0.218	0.642
IL-20	0.011	0.287	0.845
IL-30	-0.007	0.234	0.680
IL-40	-0.057	0.245	1.064
IL-50	0.000	0.215	0.972
IL-60	-0.049	0.370	1.196
IL-70	0.061	0.078	1.272
IL-80	0.042	0.081	1.150
IL-90	0.024	0.108	1.335
IN-10	-0.058	0.416	0.967
IN-20	-0.050	0.573	0.901
IN-30	0.009	0.458	1.181
IN-40	0.019	0.278	0.853
IN-50	0.054	0.524	0.786
IN-60	0.085	0.440	0.943
IN-70	0.035	0.128	1.123
IA-10	-0.076	0.146	0.549
IA-20	0.172	-0.067	0.486
IA-30	0.092	0.138	0.567
IA-40	0.019	0.082	0.612
IA-50	0.207	0.005	0.668
IA-60	0.109	0.143	0.527
IA-70	0.152	0.152	0.572
IA-80	0.199	0.126	0.687
IA-90	0.091	0.085	0.464
MI-90	0.018	0.472	1.359
MN-40	0.058	0.320	1.125
MN-50	-0.033	0.134	0.920
MN-70	-0.154	0.234	0.640
MN-80	-0.094	-0.060	0.475
MN-90	0.033	-0.011	0.776
MO-12	0.001	0.072	0.762
MO-36	-0.048	0.581	0.983
MO-45	0.003	0.119	1.043
MO-90	-0.015	0.100	0.906
NE-60	0.051	0.407	1.154
OH-10	0.075	0.426	1.291
OH-20	-0.063	0.363	1.071
OH-40	0.023	0.477	1.066
OH-50	-0.006	0.408	1.010
OH-70	-0.017	0.311	1.110
TN-12	0.009	0.271	0.467
Mean	0.023	0.240	0.886

Table 14. Out-of-Sample Forecast Errors of U.S. Total Corn Yield, 1988 – 2005/2010

Year	Actual U.S. Yield	Thompson	R&S	pGWR	Thompson CRD	R&S CRD	Three-state	Composite of 6	Composite of 4	Trend	USDA
1988	84.6	-13.385	-0.573	-14.098	-12.803	-1.999	-13.576	-9.406	-13.466	-32.145	6.100
1989	116.3	0.012	-4.967	2.501	-1.452	-5.703	-5.053	-2.444	-0.998	2.795	3.900
1990	118.5	-8.441	-3.672	-4.358	-6.247	-5.065	0.844	-4.490	-4.551	3.157	-3.200
1991	108.6	0.456	-13.226	1.064	0.496	-15.036	6.623	-3.270	2.160	-9.279	2.500
1992	131.5	-13.587	-2.987	-7.437	-15.150	-4.982	-16.319	-10.077	-13.123	13.430	10.100
1993	100.7	-24.421	-25.554	-20.090	-19.127	-19.324	-19.763	-21.380	-20.850	-21.023	-12.400
1994	138.6	7.006	7.227	7.292	4.461	3.464	-0.662	4.798	4.524	18.401	9.600
1995	113.5	6.361	-8.193	7.698	4.334	-8.837	3.180	0.757	5.394	-11.533	-7.600
1996	127.1	0.544	-6.693	-0.058	-1.903	-9.793	-6.395	-4.050	-1.953	2.051	6.900
1997	126.7	-9.843	-5.839	-8.123	-13.500	-9.794	-12.938	-10.006	-11.101	-0.287	1.500
1998	134.4	8.873	1.501	6.699	7.835	-2.470	11.754	5.699	8.790	7.826	2.400
1999	133.8	3.345	3.831	2.957	5.225	3.586	-3.022	2.654	2.126	4.814	1.600
2000	136.9	-0.693	-5.631	-1.544	-1.714	-8.996	0.724	-2.976	-0.807	6.238	-4.900
2001	138.2	4.500	2.293	-3.443	-3.000	0.847	0.580	0.296	-0.341	5.494	4.700
2002	129.3	4.694	-0.479	2.945	1.953	-1.560	1.276	1.471	2.717	-5.420	3.900
2003	142.2	9.949	3.007	11.971	13.328	-0.248	7.808	7.636	10.764	7.077	3.700
2004	160.3	3.754	10.948	2.010	3.136	9.615	3.120	5.430	3.005	23.271	10.900
2005	147.9	8.621	11.411	6.230	6.549	8.274	5.208	7.715	6.652	7.295	4.700
2006	149.1	9.522	-	7.182	7.095	-	1.806	-	6.402	6.738	-5.600
2007	150.7	11.647	-	7.167	8.132	-	0.599	-	6.886	6.486	-5.100
2008	153.9	4.668	-	2.241	-0.403	-	-0.640	-	1.467	7.565	1.600
2009	164.7	1.939	-	-2.004	-3.972	-	-12.379	-	-4.104	16.614	2.800
2010	152.8	6.734	-	6.643	11.241	-	14.841	-	9.864	1.888	-9.700
Mean	-	0.968	-2.089	0.584	-0.239	-3.779	-1.408	-1.758	-0.024	2.672	1.235
Standard deviation	-	9.081	8.773	7.608	8.574	7.575	8.815	7.540	8.132	12.291	6.308

Note: “Thompson,” “R&S,” “pGWR,” “Thompson CRD,” “R&S CRD,” “Three-state,” “Composite of 6,” “Composite of 4,” “Trend,” and “USDA” denote forecast errors from modified Thompson corn model, modified Roberts and Schlenker model, modified Thompson CRD level corn model, modified Roberts and Schlenker CRD level model, three-state model, average of “Thompson,” “R&S,” “pGWR,” “Thompson CRD,” “R&S CRD,” and “Three-state” models, average of “Thompson,” “pGWR,” “Thompson CRD,” and “Three-state” models, trend only model, and USDA forecast, respectively.

Table 15. Out-of-Sample Forecast Accuracy Statistics for Corn Yields

		Recursive method, 1988-2005									**Recursive method, 1988-2010**						
	Model	Thompson	R&S	pGWR	Thompson CRD	R&S CRD	Three-state	Composite of 6	Trend	USDA	Thompson	pGWR	Thompson CRD	Three-state	Composite of 4	Trend	USDA
Area-weighted forecast	**RMSE**	9.245	8.778	7.906	8.692	8.275	8.749	7.535	13.045	6.449	8.934	7.464	8.389	8.736	7.954	12.314	6.292
	RMSPE	8.419	7.796	7.357	7.681	7.192	7.889	6.837	12.276	5.330	7.820	6.733	7.143	7.439	7.018	11.188	5.027
	MAE	7.138	6.557	6.140	6.790	6.644	6.602	5.809	10.085	5.589	7.087	5.902	6.655	6.483	6.176	9.601	5.452
	MAPE	6.026	5.387	5.253	5.680	5.487	5.648	4.881	8.586	4.520	5.705	4.833	5.323	5.257	5.016	7.809	4.243
Average CRD forecast	**RMSE**	16.102	15.742	17.420	19.487	15.974	-	-	-	-	15.818	17.147	19.188	-	-	-	-
	RMSPE	18.334	17.719	19.191	20.456	17.708	-	-	-	-	16.818	17.696	18.901	-	-	-	-
	MAE	12.011	11.771	12.807	14.545	12.187	-	-	-	-	12.028	12.841	14.411	-	-	-	-
	MAPE	11.169	10.912	11.809	13.192	11.222	-	-	-	-	10.422	11.065	12.277	-	-	-	-

		6-year resampling method									**7-year resampling method**						
	Model	Thompson	R&S	pGWR	Thompson CRD	R&S CRD	Three-state	Composite of 6	Trend	USDA	Thompson	pGWR	Thompson CRD	Three-state	Composite of 4	Trend	USDA
Area-weighted forecast	**RMSE**	6.579	7.401	6.533	6.925	7.352	6.930	5.914	10.517	4.978	6.471	6.347	6.625	6.688	6.164	9.960	5.053
	RMSPE	6.894	7.864	7.079	7.252	7.805	7.265	6.106	10.923	4.919	6.653	6.817	6.866	6.757	6.446	10.164	4.816
	MAE	5.284	5.945	5.275	5.593	5.970	5.530	4.731	8.560	4.052	5.015	5.039	5.266	5.319	4.881	7.957	4.141
	MAPE	5.484	6.216	5.579	5.794	6.206	5.700	4.830	8.780	4.063	5.084	5.248	5.357	6.688	6.164	9.960	5.053
Average CRD forecast	**RMSE**	12.642	12.995	14.291	16.142	13.299	-	-	16.949	-	12.583	14.264	15.461	-	-	16.498	-
	RMSPE	15.920	15.847	17.937	20.104	16.136	-	-	22.178	-	15.254	17.371	18.487	-	-	20.914	-
	MAE	9.508	9.969	10.689	11.915	10.275	-	-	12.912	-	9.509	10.724	11.484	-	-	12.580	-
	MAPE	10.571	10.943	11.882	13.210	11.255	-	-	14.491	-	10.110	11.429	12.125	-	-	13.534	-

Note: "Thompson," "R&S," "pGWR," "Thompson CRD," "R&S CRD," "Three-state," "Composite of 6," "Composite of 4," "Trend," and "USDA" denote modified Thompson corn model, modified Roberts and Schlenker model, modified Thompson CRD level corn model, modified Roberts and Schlenker CRD level model, three-state model, average of "Thompson," "R&S," "pGWR," "Thompson CRD," "R&S CRD," and "Three-state" models, average of "Thompson," "pGWR," "Thompson CRD," and "Three-state" models, trend only model, and USDA forecast, respectively.

Table 16. Out-of-Sample Forecast Errors of U.S. Total Soybean Yield, 1988 – 2005/2010

Year	Actual U.S. Yield	Thompson	R&S	pGWR	Thompson CRD	R&S CRD	Three-state	Composite of 6	Composite of 4	Trend	USDA
1988	27.0	-4.091	-1.244	-3.986	-3.115	-1.166	-0.120	-2.287	-2.828	-6.225	1.100
1989	32.3	0.001	-0.762	-0.122	-0.856	-0.845	-0.335	-0.487	-0.328	-0.095	0.300
1990	34.1	2.403	0.842	2.738	2.745	1.402	5.333	2.577	3.305	1.140	1.700
1991	34.2	2.455	-2.091	2.772	3.209	-1.855	4.020	1.418	3.114	0.469	3.200
1992	37.6	0.694	3.014	0.710	1.062	2.770	2.621	1.812	1.272	3.345	1.700
1993	32.6	-1.125	-1.950	-2.475	-1.714	-0.312	-4.277	-1.975	-2.398	-2.507	-1.400
1994	41.4	4.270	4.789	4.286	4.211	4.194	4.052	4.300	4.205	6.236	3.200
1995	35.3	1.649	-1.010	1.229	1.095	-0.745	1.597	0.636	1.392	-0.980	-1.700
1996	37.6	1.321	-0.040	1.182	1.079	-0.479	1.152	0.702	1.183	0.690	1.800
1997	38.9	-0.035	1.913	-0.094	-0.104	1.046	-1.548	0.196	-0.445	1.620	-0.400
1998	38.9	0.136	-0.320	0.249	1.128	-0.645	1.290	0.306	0.701	1.312	-1.700
1999	36.6	-1.682	-2.068	-1.529	-1.994	-1.957	-2.145	-1.896	-1.838	-1.659	-1.300
2000	38.1	-2.637	-2.731	-3.217	-2.533	-3.482	-3.300	-2.983	-2.922	-0.460	-1.400
2001	39.6	0.149	-0.665	-0.535	-0.773	-0.954	-0.615	-0.565	-0.443	0.572	1.400
2002	38.0	-0.759	-1.261	-0.749	-0.448	-1.248	-1.536	-1.000	-0.873	-1.313	1.000
2003	33.9	-5.659	-6.398	-6.290	-5.739	-6.950	-6.373	-6.235	-6.015	-5.977	-2.500
2004	42.2	-1.333	0.977	-1.806	-0.799	1.019	-3.012	-0.826	-1.738	2.753	3.700
2005	43.1	2.843	3.510	2.960	3.374	3.071	2.905	3.110	3.021	2.884	3.500
2006	42.9	1.640	-	1.472	1.774	-	1.093	-	1.495	2.253	1.100
2007	41.7	-0.064	-	0.216	1.086	-	-1.885	-	-0.162	0.498	0.300
2008	39.7	-1.088	-	-1.352	-1.083	-	-1.988	-	-1.378	-1.968	-0.300
2009	44.0	-0.288	-	-0.872	-0.896	-	-1.936	-	-0.998	2.122	1.700
2010	43.5	1.365	-	1.332	1.626	-	1.619	-	1.486	1.062	-1.200
Mean	-	0.007	-0.305	-0.169	0.102	-0.396	-0.147	-0.178	-0.052	0.251	0.600
Standard deviation	-	2.247	2.591	2.432	2.322	2.538	2.901	2.465	2.396	2.800	1.846

Note: “Thompson,” “R&S,” “pGWR,” “Thompson CRD,” “R&S CRD,” “Three-state,” “Composite of 6,” “Composite of 4,” “Trend,” and “USDA” denote forecast errors from modified Thompson corn model, modified Roberts and Schlenker model, modified Thompson CRD level corn model, modified Roberts and Schlenker CRD level model, three-state model, average of “Thompson,” “R&S,” “pGWR,” “Thompson CRD,” “R&S CRD,” and “Three-state” models, average of “Thompson,” “pGWR,” “Thompson CRD,” and “Three-state” models, trend only model, and USDA forecast, respectively.

Table 17. Out-of-Sample Forecast Accuracy Statistics for Soybean Yields

		Recursive method, 1988-2005									Recursive method, 1988-2010						
	Model	Thompson	R&S	pGWR	Thompson CRD	R&S CRD	Three-state	Composite of 6	Trend	USDA	Thompson	pGWR	Thompson CRD	Three-state	Composite of 4	Trend	USDA
Area-weighted forecast	**RMSE**	2.418	2.536	2.627	2.470	2.498	3.079	2.402	2.973	2.075	2.198	2.384	2.273	2.841	2.344	2.750	1.902
	RMSPE	7.040	6.921	7.614	7.027	6.878	8.613	6.729	8.768	5.465	6.340	6.854	6.387	7.859	6.590	7.985	4.968
	MAE	1.847	1.977	2.052	1.999	1.897	2.569	1.851	2.235	1.833	1.639	1.834	1.845	2.381	1.893	2.093	1.635
	MAPE	5.220	5.387	5.792	5.605	5.160	7.024	5.141	6.297	4.919	4.542	5.072	5.049	6.375	5.184	5.740	4.313
Average CRD forecast	**RMSE**	4.822	4.897	5.120	5.144	4.940	-	-	-	-	4.690	4.968	5.037	-	-	-	-
	RMSPE	15.172	15.047	16.401	15.410	14.619	-	-	-	-	14.222	15.343	14.486	-	-	-	-
	MAE	3.714	3.867	3.821	3.874	3.806	-	-	-	-	3.582	3.716	3.794	-	-	-	-
	MAPE	10.411	10.642	10.743	10.624	10.419	-	-	-	-	9.674	10.058	10.008	-	-	-	-

		6-year resampling method									7-year resampling method						
	Model	Thompson	R&S	pGWR	Thompson CRD	R&S CRD	Three-state	Composite of 6	Trend	USDA	Thompson	pGWR	Thompson CRD	Three-state	Composite of 4	Trend	USDA
Area-weighted forecast	**RMSE**	1.700	1.854	1.837	1.724	1.842	2.066	1.641	2.232	1.553	1.691	1.850	1.758	2.074	1.784	2.225	1.512
	RMSPE	5.621	5.977	6.082	5.651	5.965	6.696	5.314	7.300	4.574	5.536	6.070	5.712	6.577	5.778	7.242	4.412
	MAE	1.357	1.463	1.457	1.360	1.460	1.645	1.307	1.735	1.239	1.323	1.438	1.360	1.657	1.381	1.738	1.208
	MAPE	4.477	4.757	4.817	4.461	4.754	5.346	4.262	5.623	3.800	4.281	4.669	4.373	5.237	4.435	5.546	3.637
Average CRD forecast	**RMSE**	3.640	3.896	3.857	3.936	3.916	-	-	4.441	-	3.702	3.952	3.948	-	-	4.501	-
	RMSPE	12.233	13.275	13.135	12.861	13.064	-	-	15.515	-	12.270	13.224	12.644	-	-	15.568	-
	MAE	2.816	3.054	2.958	2.988	3.037	-	-	3.466	-	2.840	2.990	2.985	-	-	3.497	-
	MAPE	8.878	9.646	9.352	9.321	9.525	-	-	11.029	-	8.758	9.236	9.080	-	-	10.921	-

Note: "Thompson," "R&S," "pGWR," "Thompson CRD," "R&S CRD," "Three-state," "Composite of 6," "Composite of 4," "Trend," and "USDA" denote modified Thompson corn model, modified Roberts and Schlenker model, modified Thompson CRD level corn model, modified Roberts and Schlenker CRD level model, three-state model, average of "Thompson," "R&S," "pGWR," "Thompson CRD," "R&S CRD," and "Three-state" models, average of "Thompson," "pGWR," "Thompson CRD," and "Three-state" models, trend only model, and USDA forecast, respectively.

Figure 1. Major Corn Production CRDs

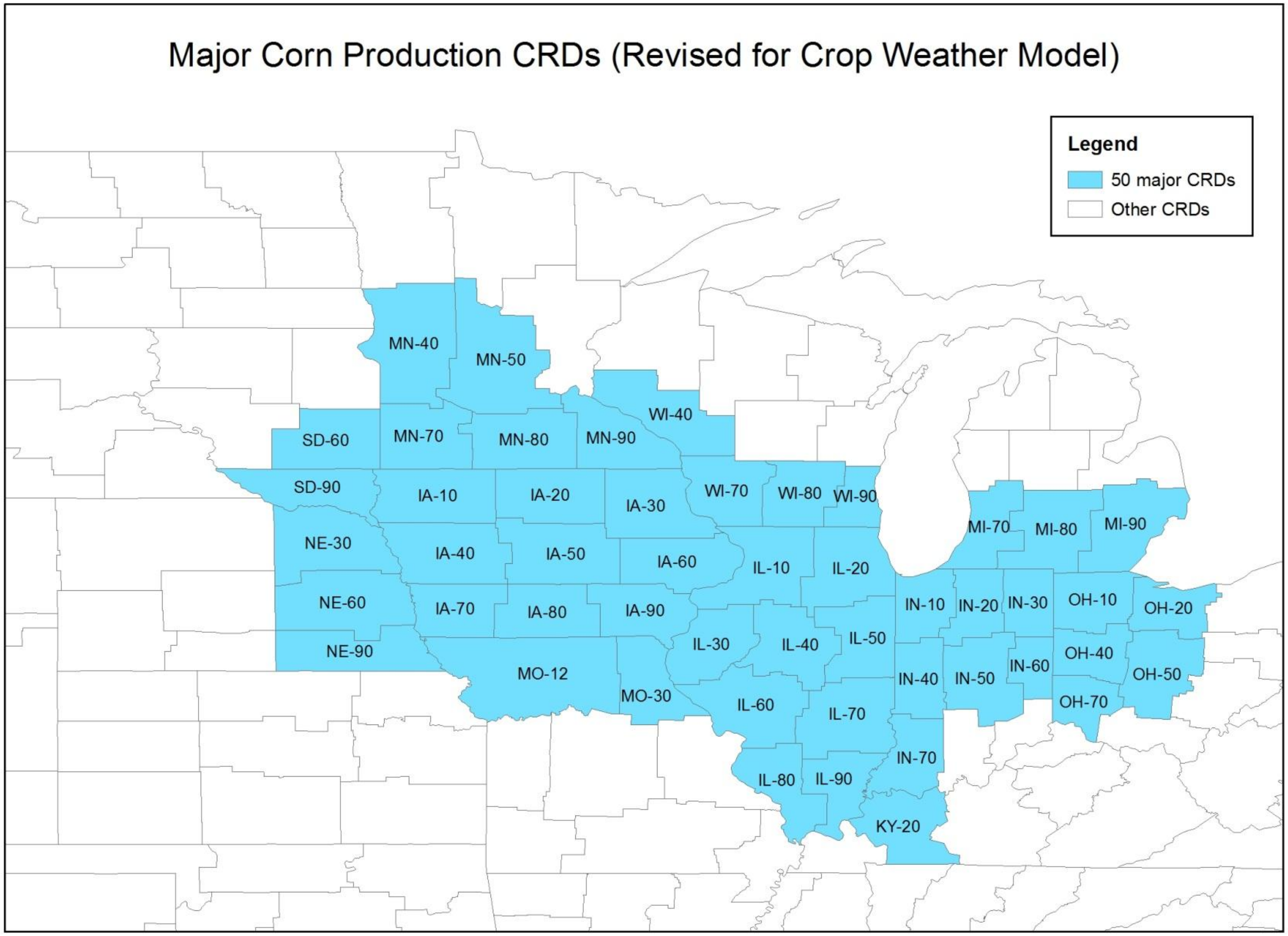


Figure 2. Major Soybean Production CRDs

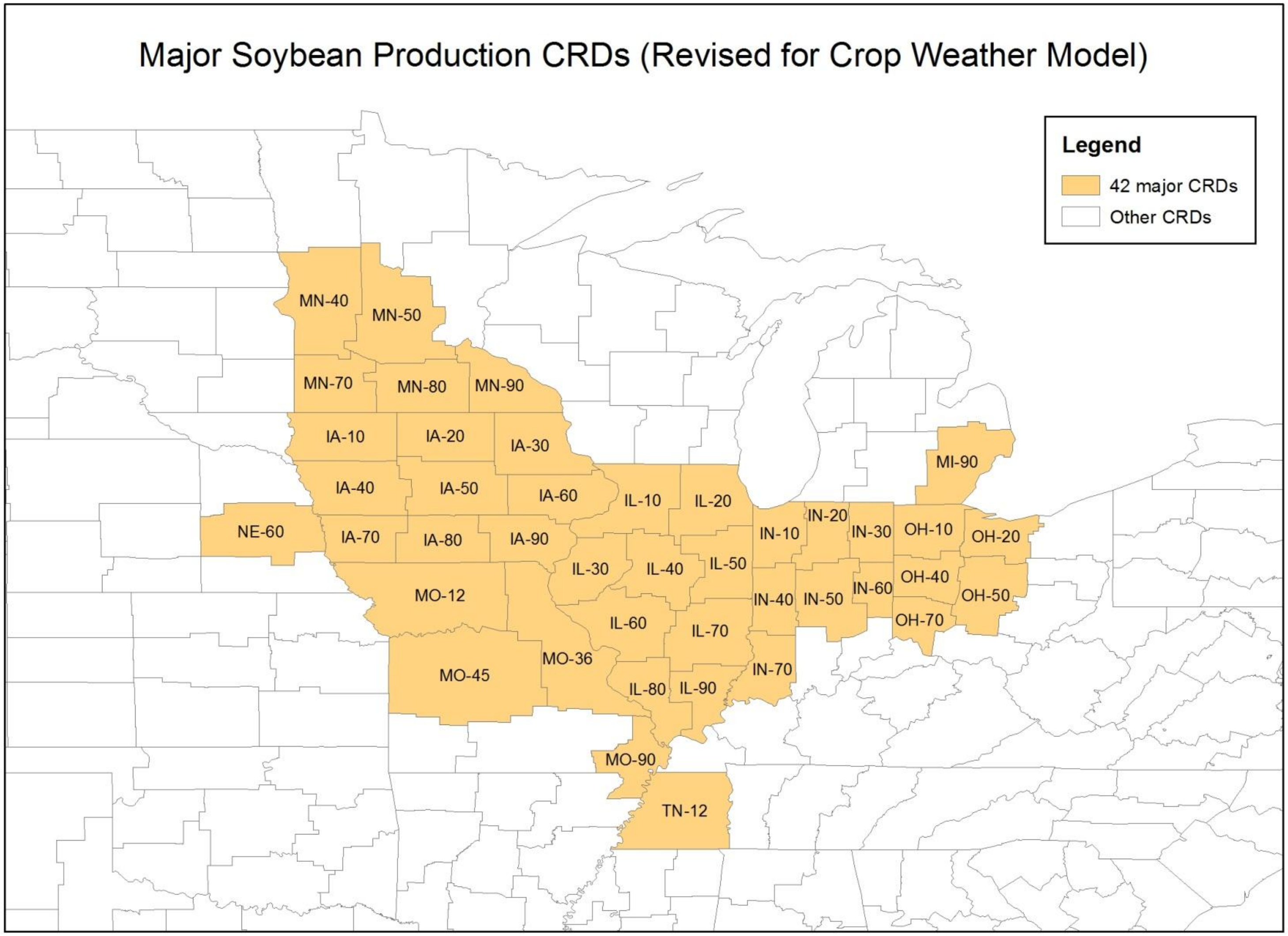
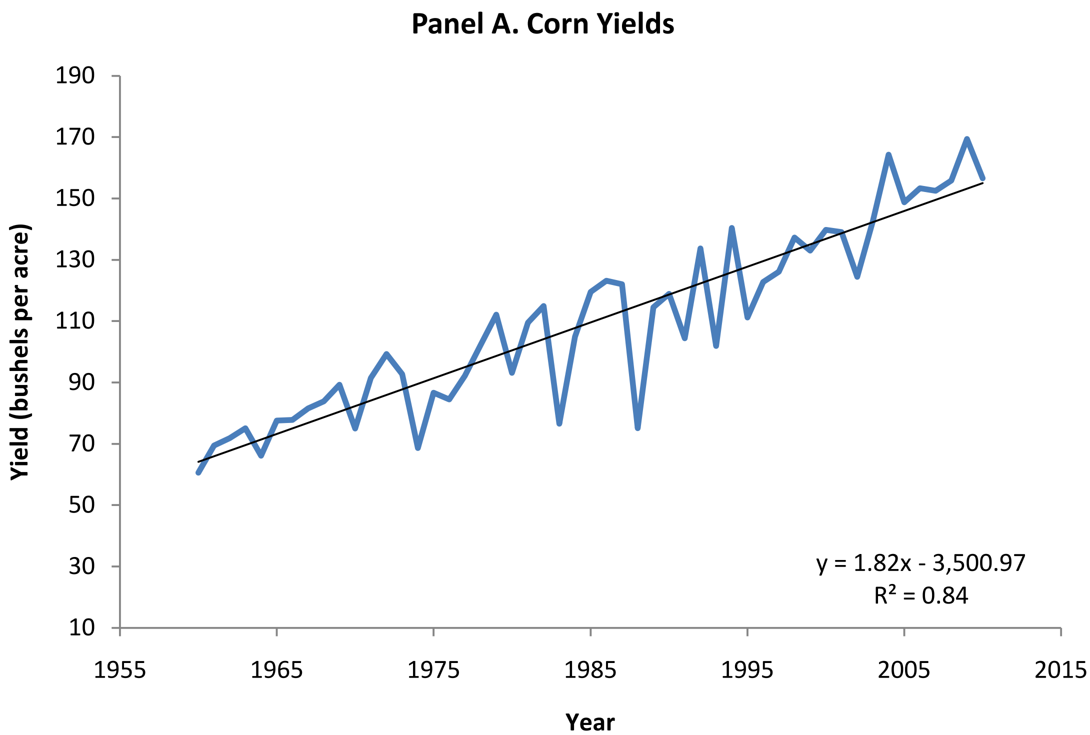


Figure 3. Average Crop Yields of Sample CRDs, 1960 – 2010



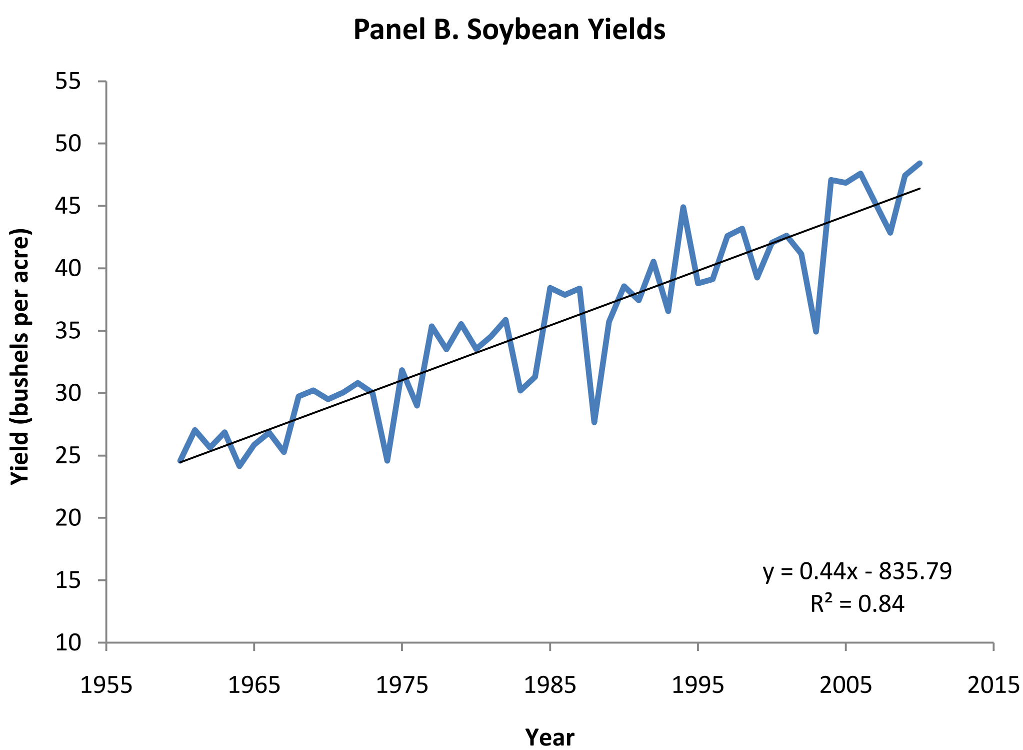
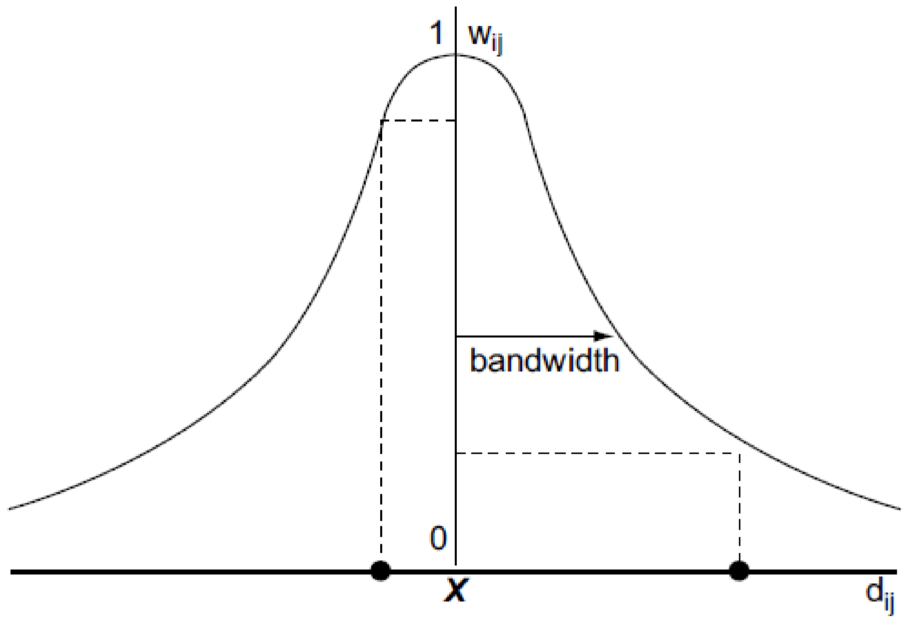
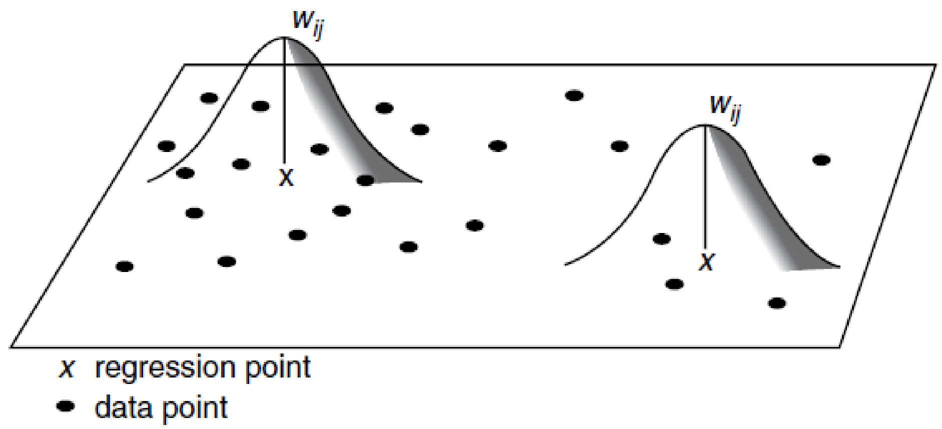


Figure 4. A Spatial Kernel



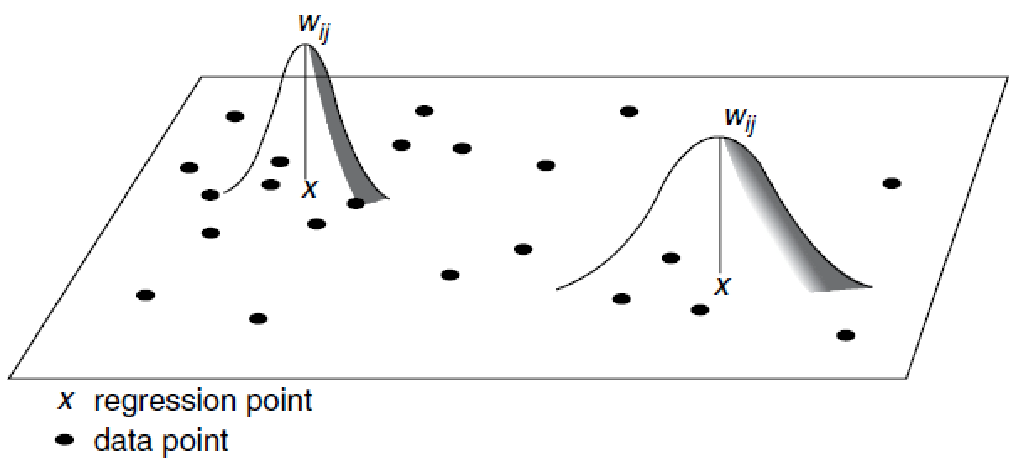
Source: Geographically Weighted Regression (2002)

Figure 5. GWR with Fixed Spatial Kernel



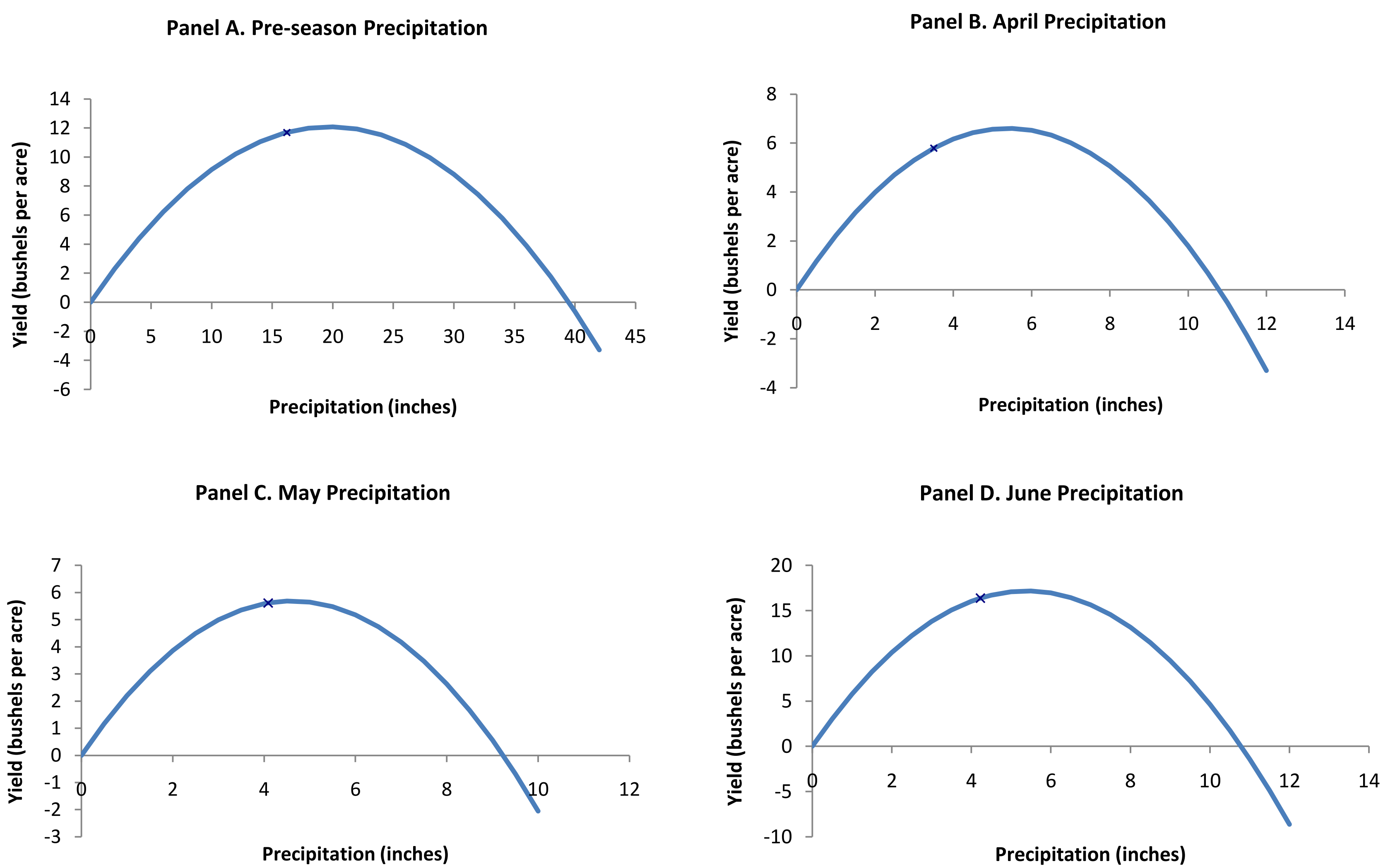
Source: Geographically Weighted Regression (2002)

Figure 6. GWR with Adaptive Spatial Kernel



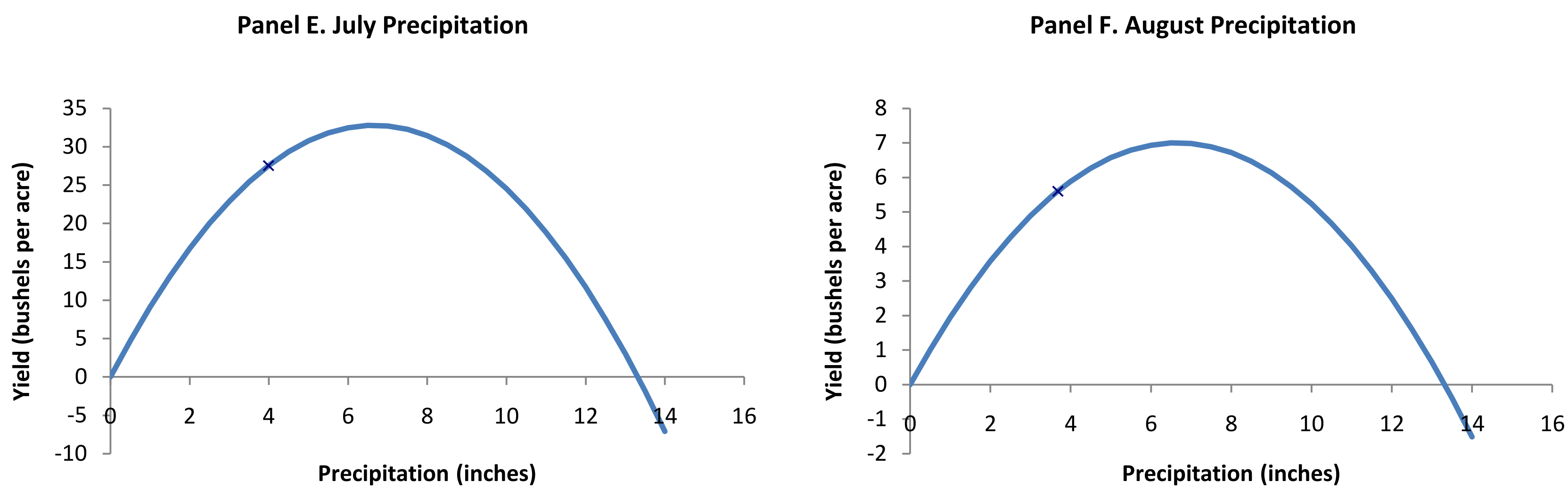
Source: Geographically Weighted Regression (2002)

Figure 7. Response of Corn Yields to Precipitation in Modified Thompson Model, 1960 – 2010



Note: “x” denotes average precipitation over the period of 1960 through 2010.

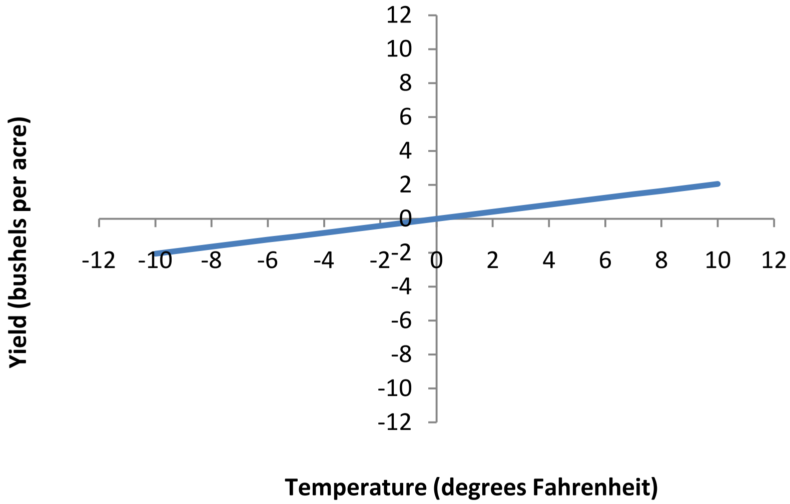
Figure 7. Response of Corn Yields to Precipitation in Modified Thompson Model, 1960 – 2010 (Continued)



Note: "x" denotes average precipitation over the period of 1960 through 2010.

Figure 8. Response of Corn Yields to Temperature in Modified Thompson Model, 1960 – 2010

Panel A. Pre-silking Temperature



Panel B. Post-silking Temperature

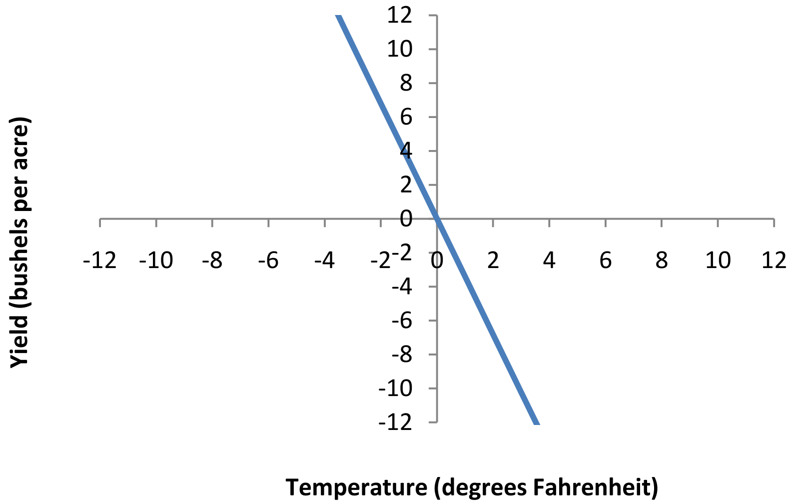
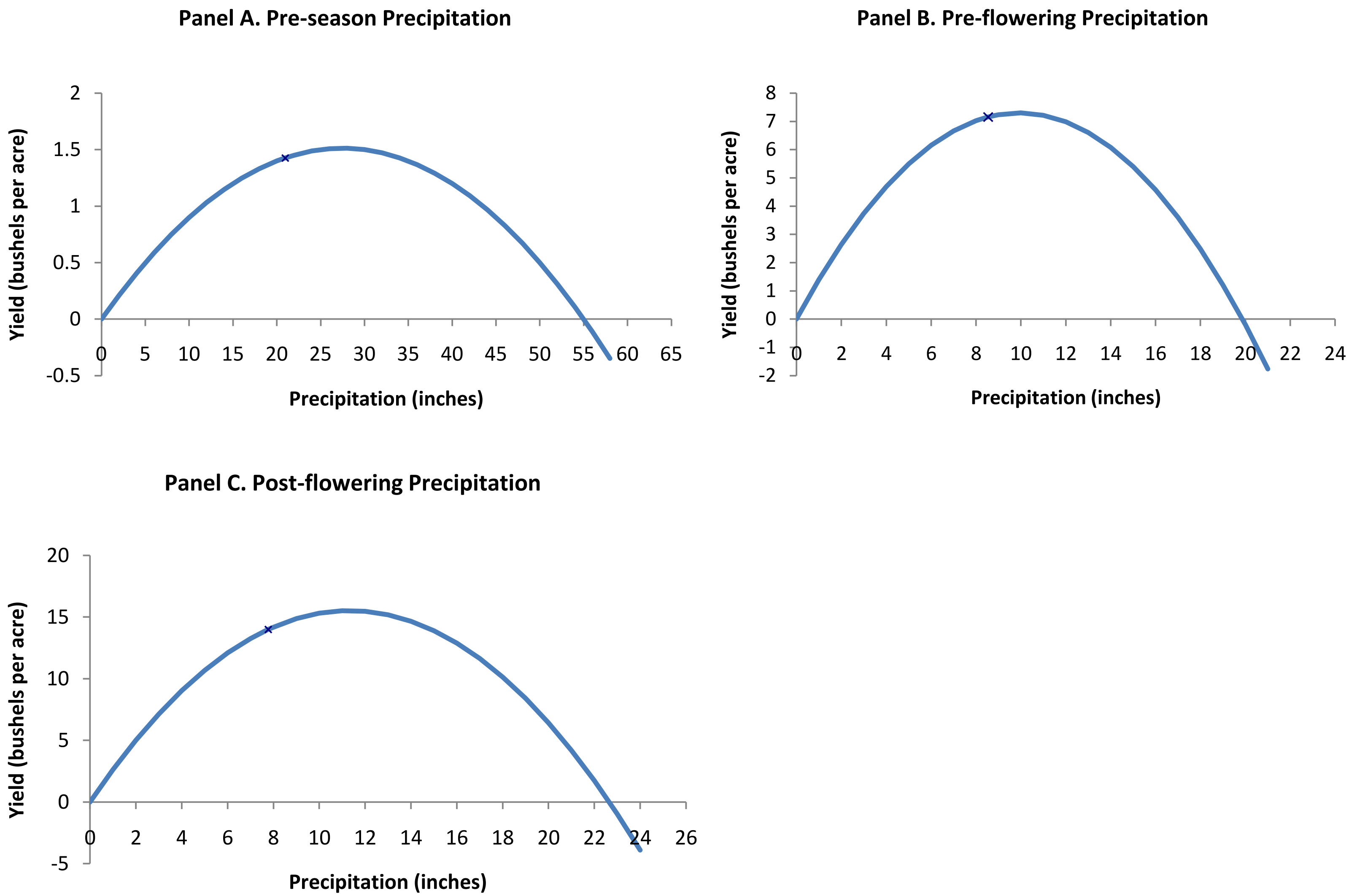


Figure 9. Response of Soybean Yields to Precipitation in Modified Thompson Model, 1960 – 2010



Note: "x" denotes average precipitation over the period of 1960 through 2010.

Figure 10. Response of Soybean Yields to Temperature in Modified Thompson Model, 1960 – 2010

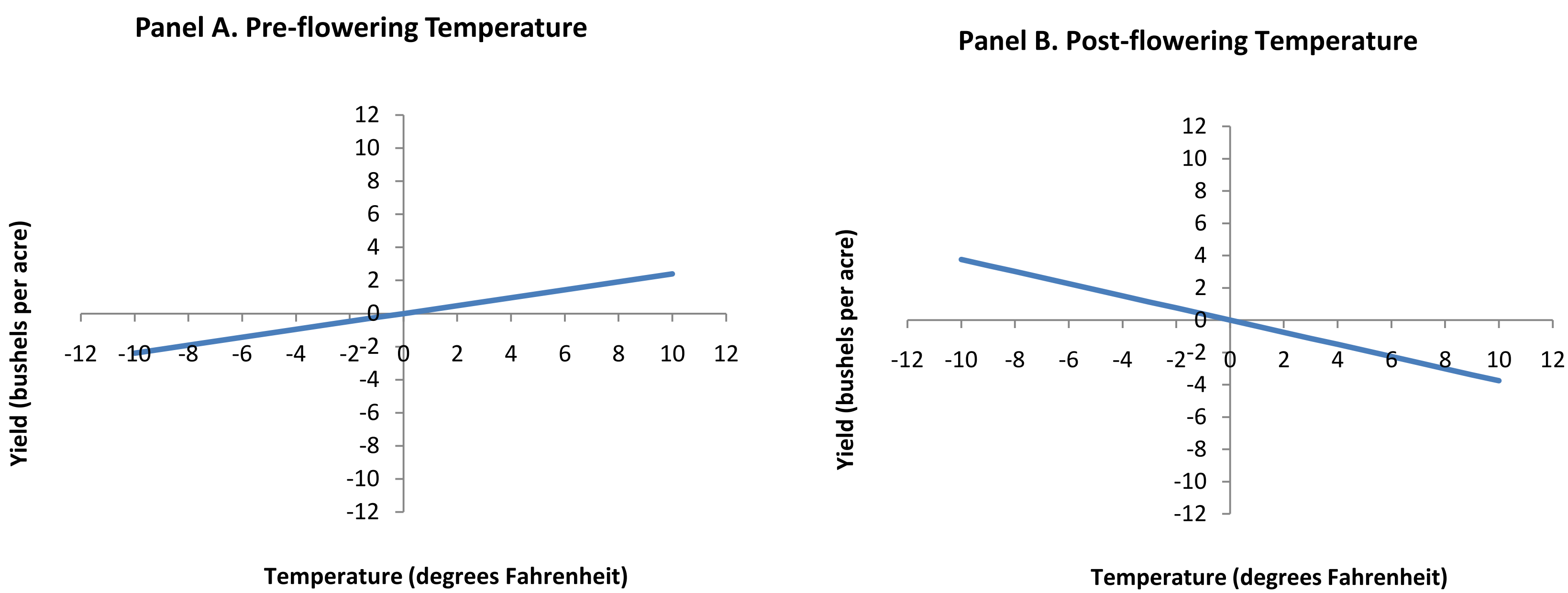
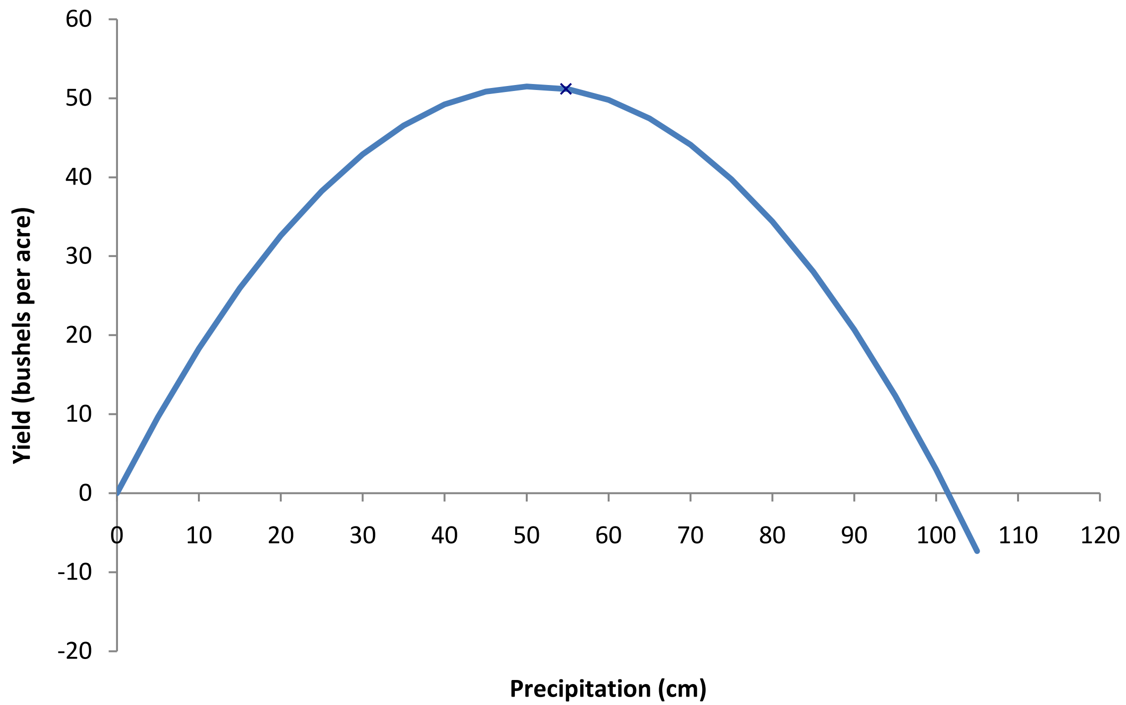


Figure 11. Response of Corn Yields to Precipitation in Modified Roberts and Schlenker Model, 1960 – 2005



Note: "x" denotes average precipitation over the period of 1960 through 2005.

Figure 12. Response of Corn Yields to Temperature in Modified Roberts and Schlenker Model, 1960 – 2005

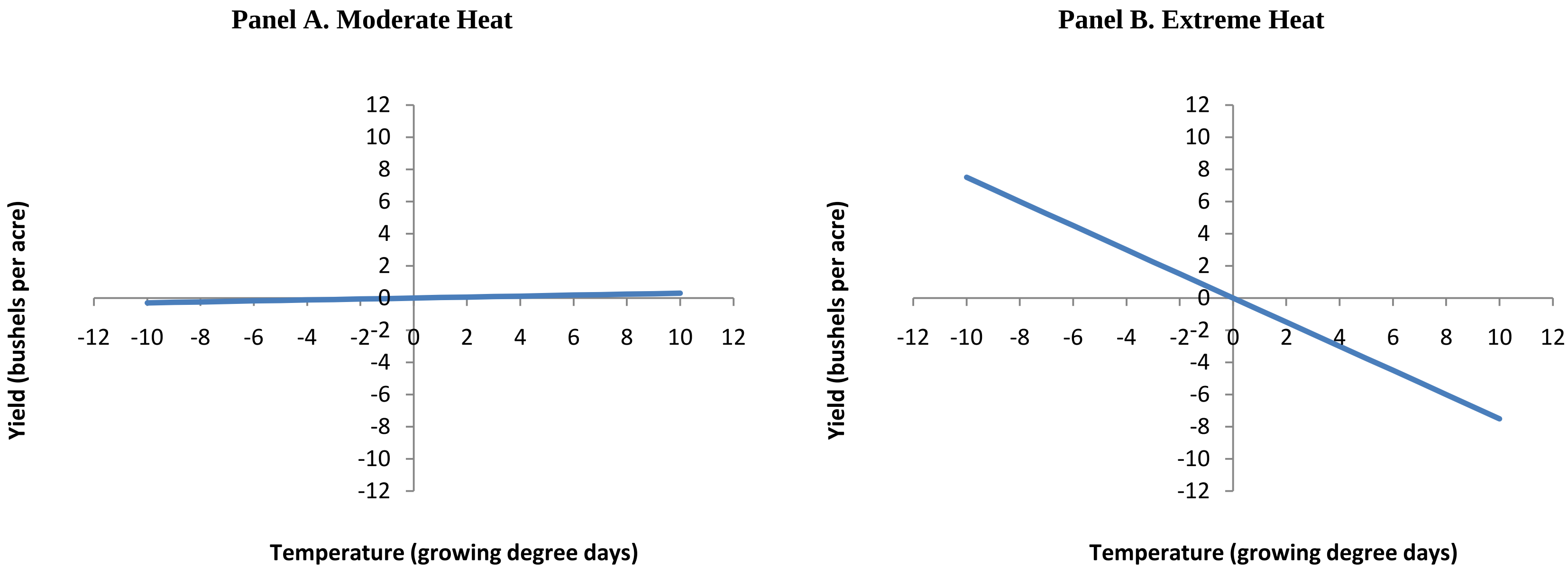
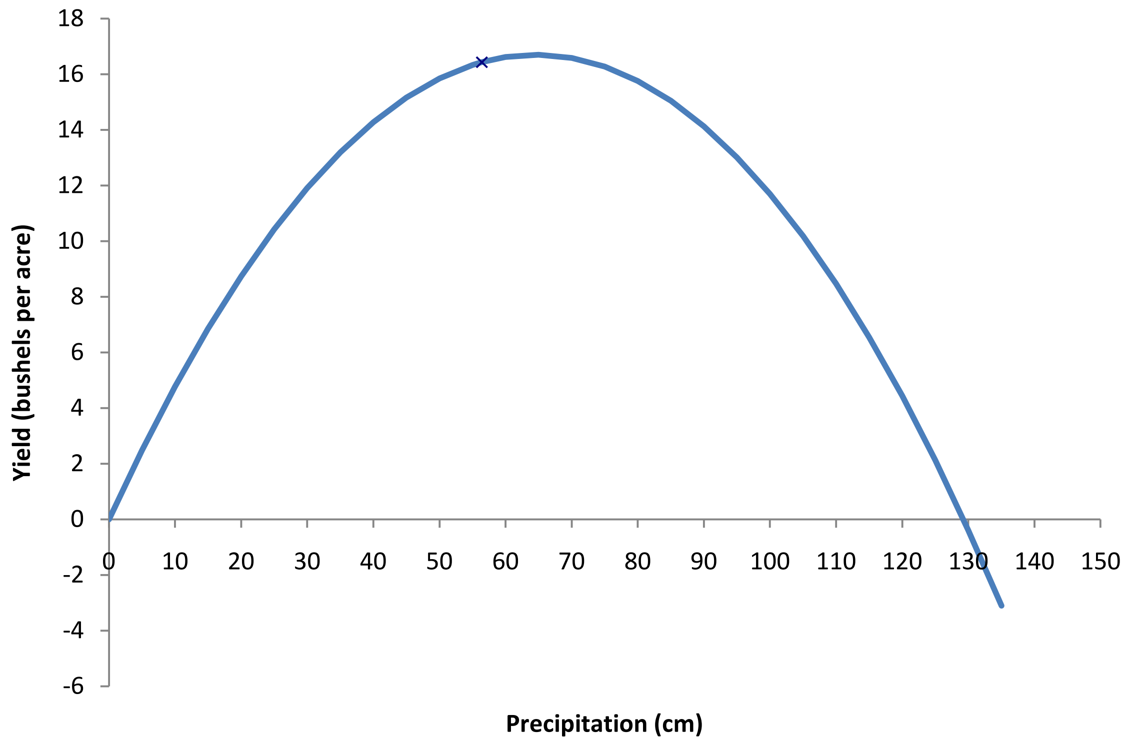


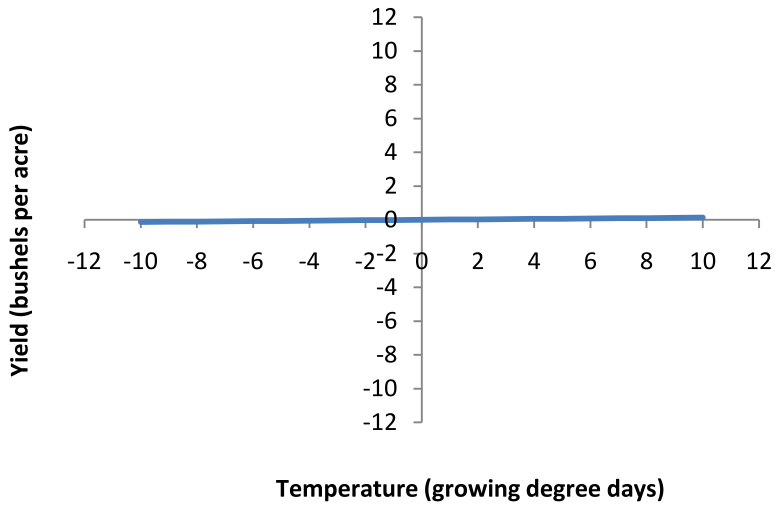
Figure 13. Response of Soybean Yields to Precipitation in Modified Roberts and Schlenker Model, 1960 – 2005



Note: “x” denotes average precipitation over the period of 1960 through 2005.

Figure 14. Response of Soybean Yields to Temperature in Modified Roberts and Schlenker Model, 1960 – 2005

Panel A. Moderate Heat



Panel B. Extreme Heat

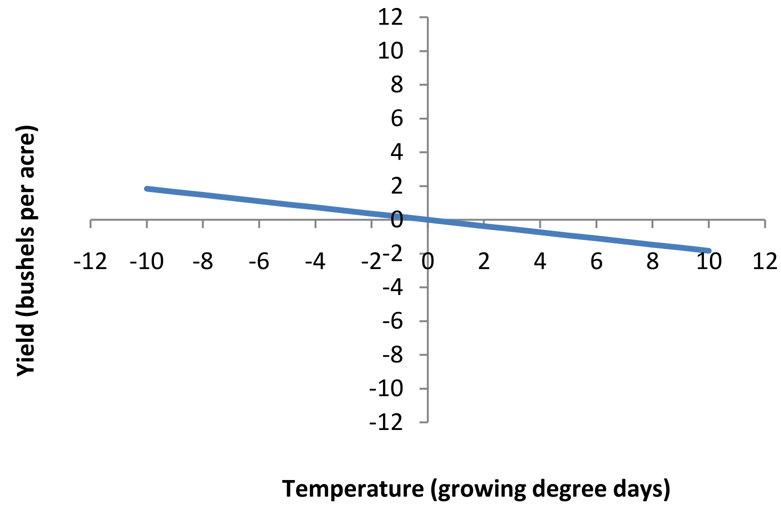


Figure 15. Coefficient Estimate of Trend in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

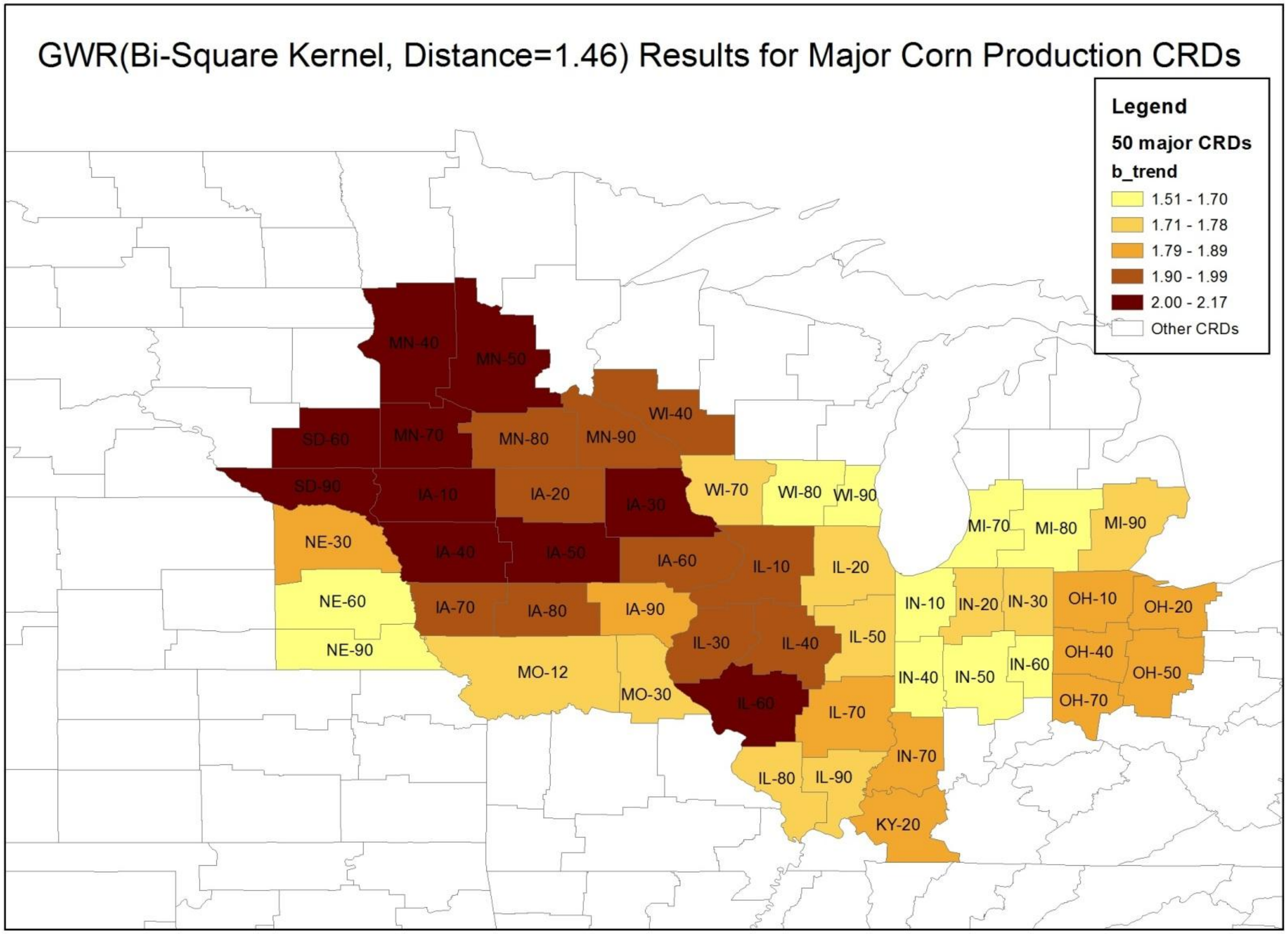


Figure 16. Coefficient Estimate of Late Planting in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

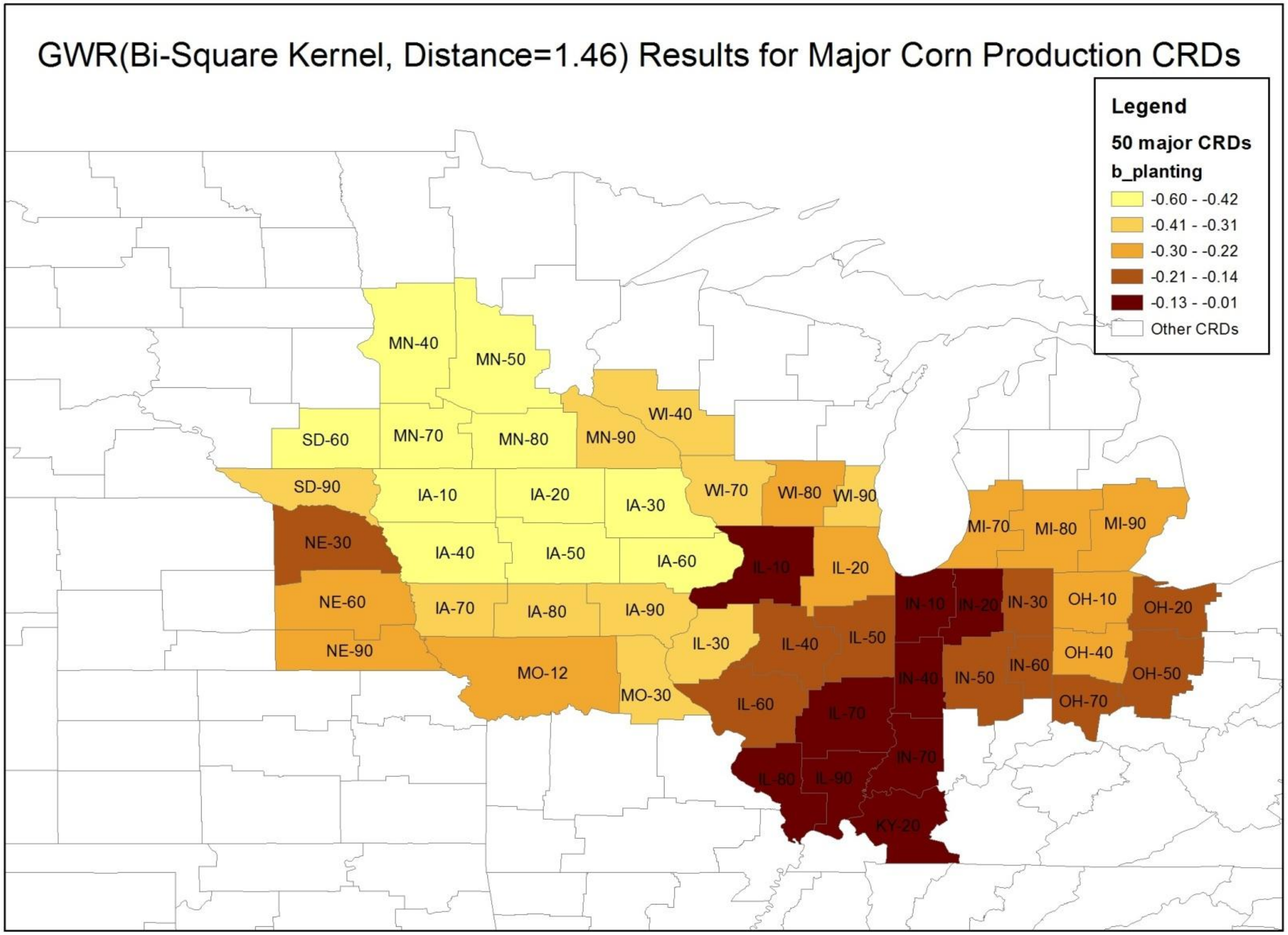


Figure 17. Marginal Effect of Pre-season Precipitation in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

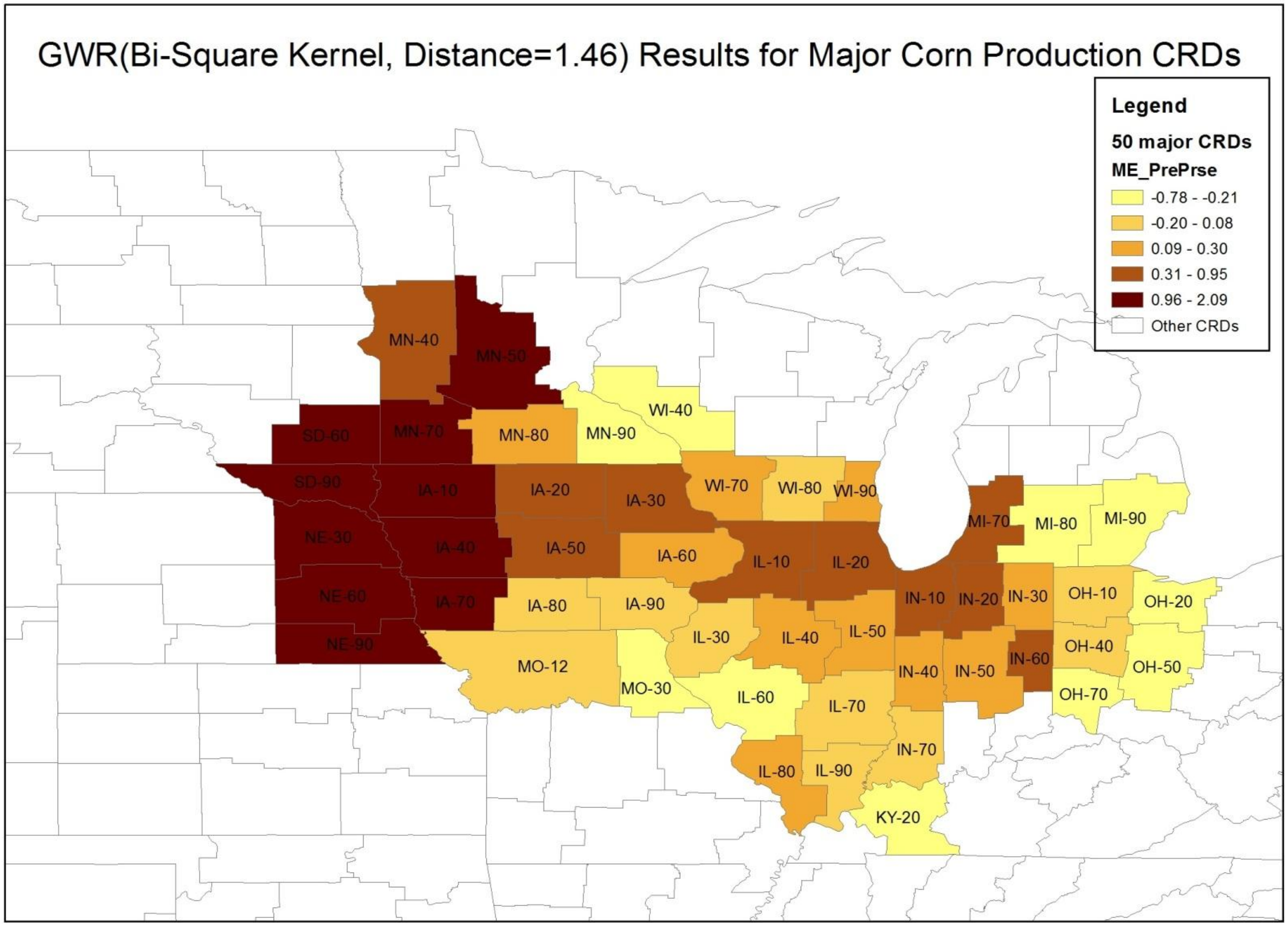


Figure 18. Marginal Effect of April Precipitation in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

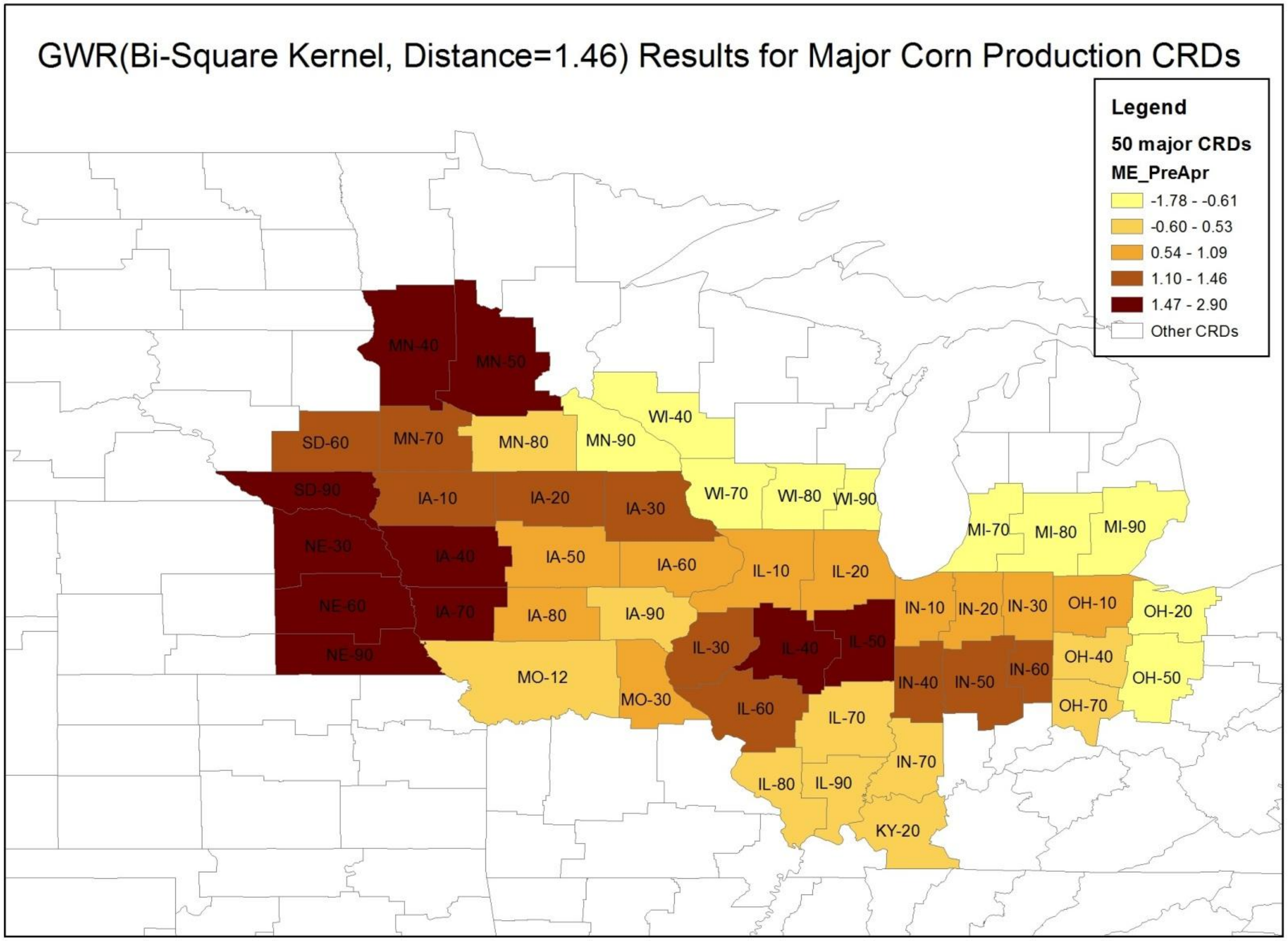


Figure 19. Marginal Effect of May Precipitation in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

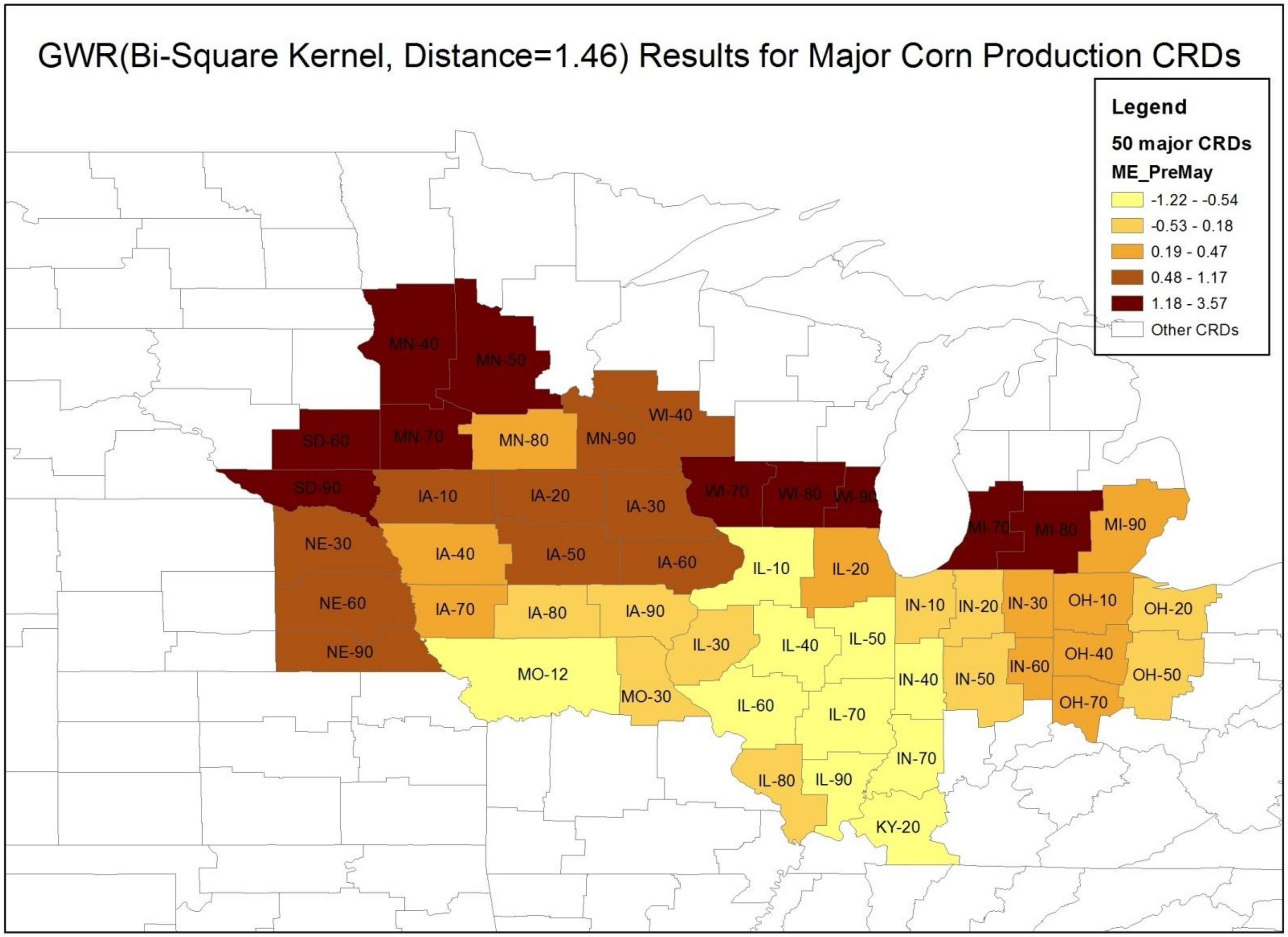


Figure 20. Marginal Effect of June Precipitation in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

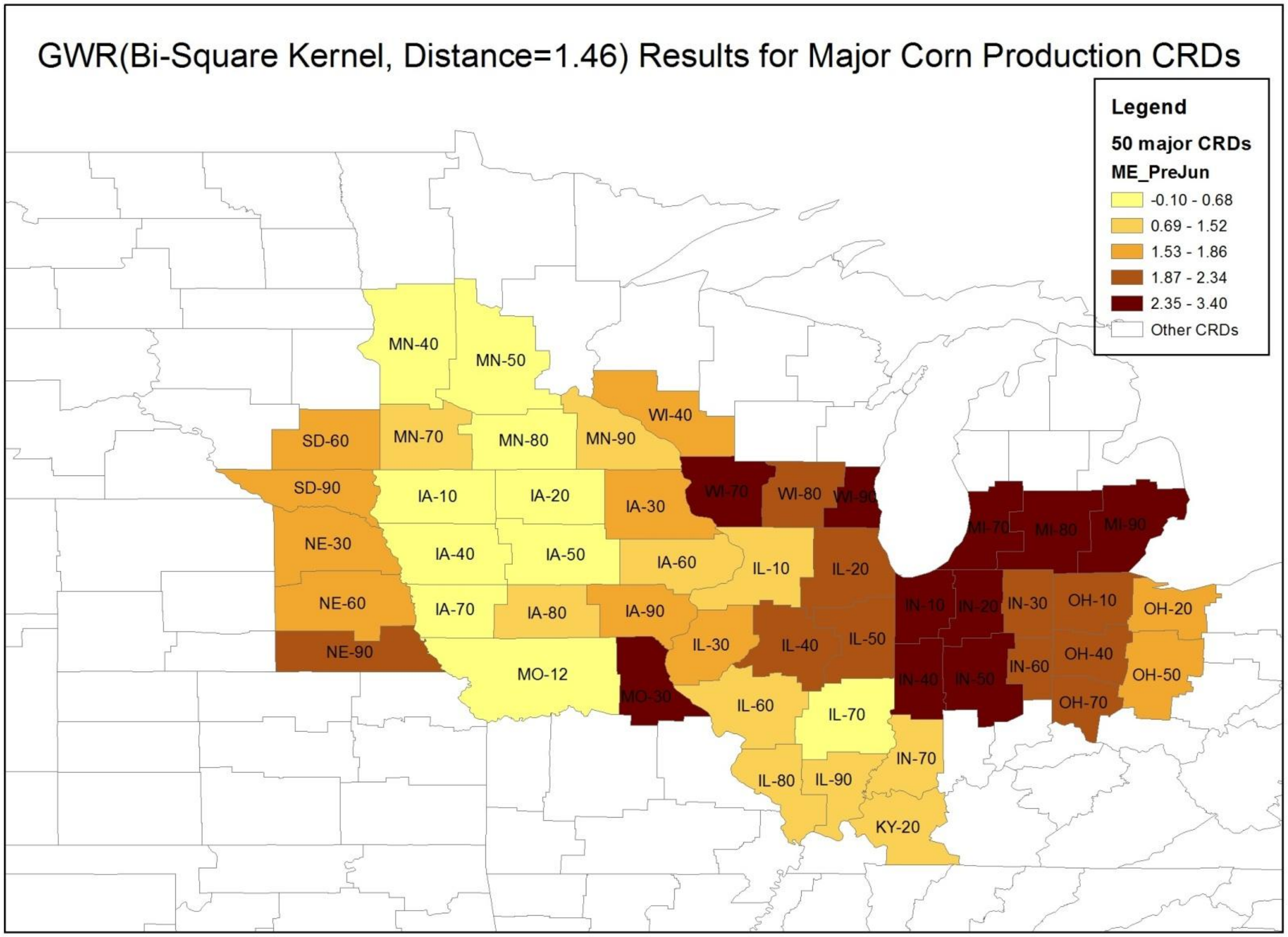


Figure 21. Marginal Effect of July Precipitation in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

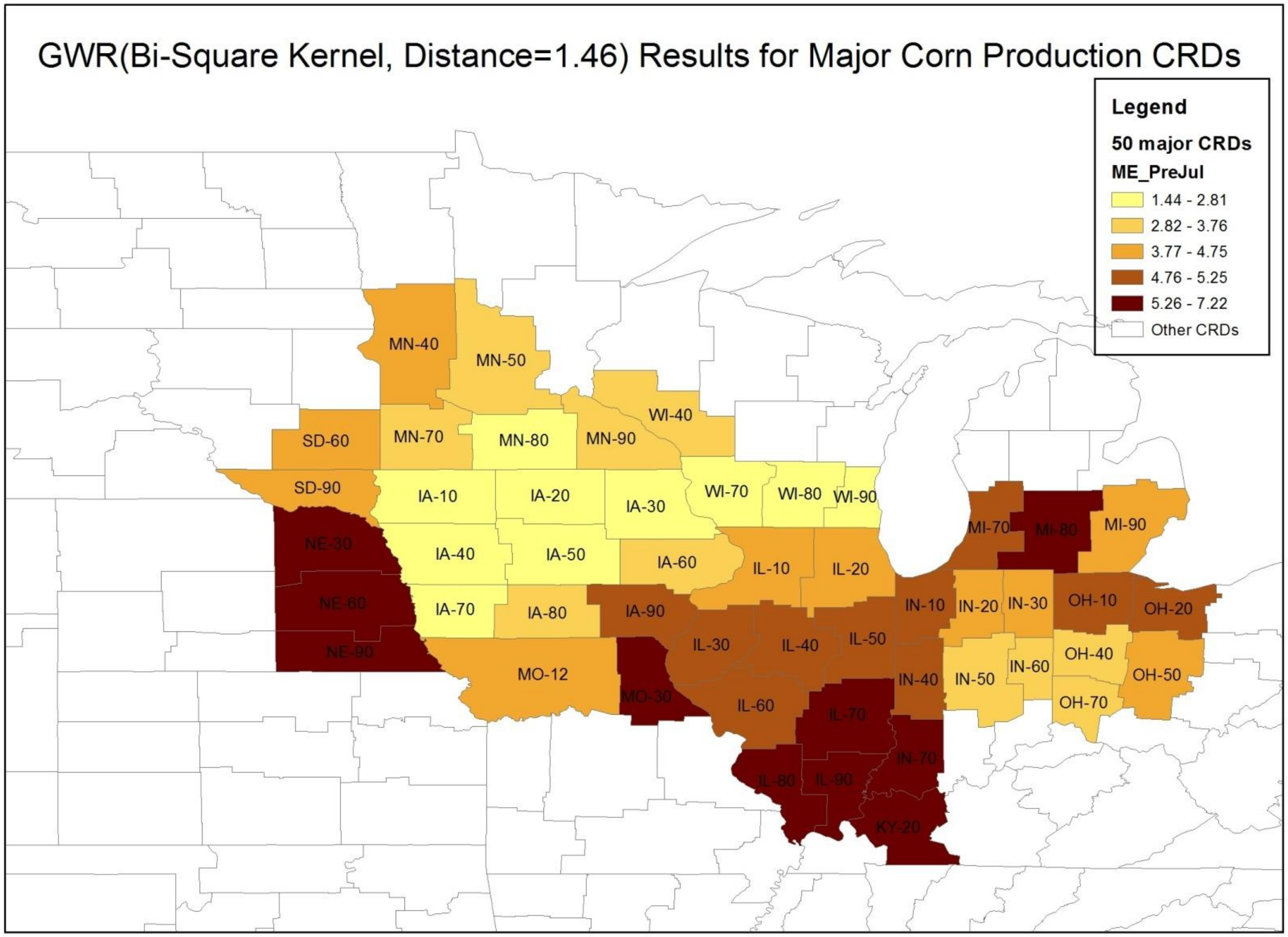


Figure 22. Marginal Effect of August Precipitation in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

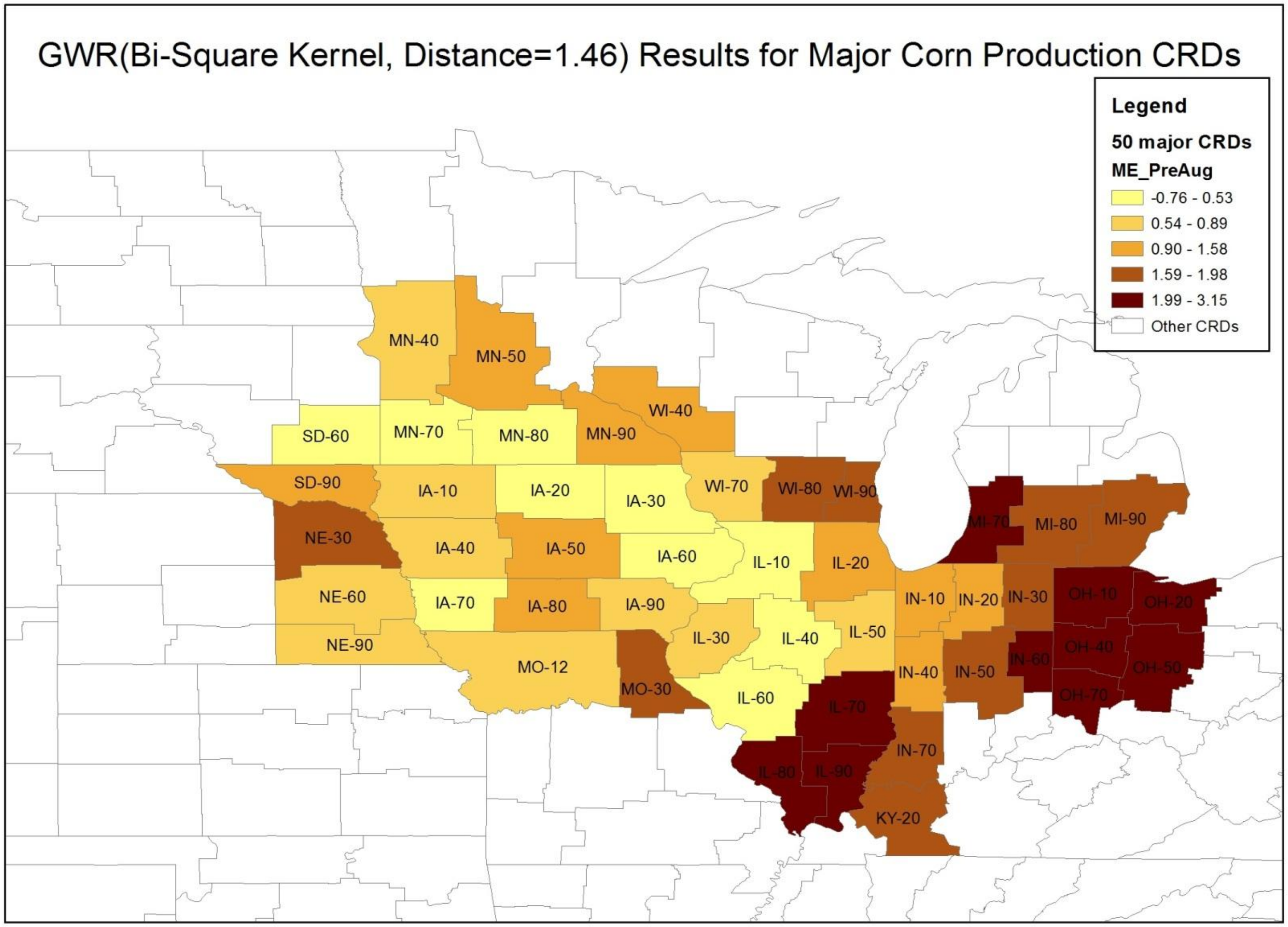


Figure 23. Coefficient Estimate of Pre-silking Temperature in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

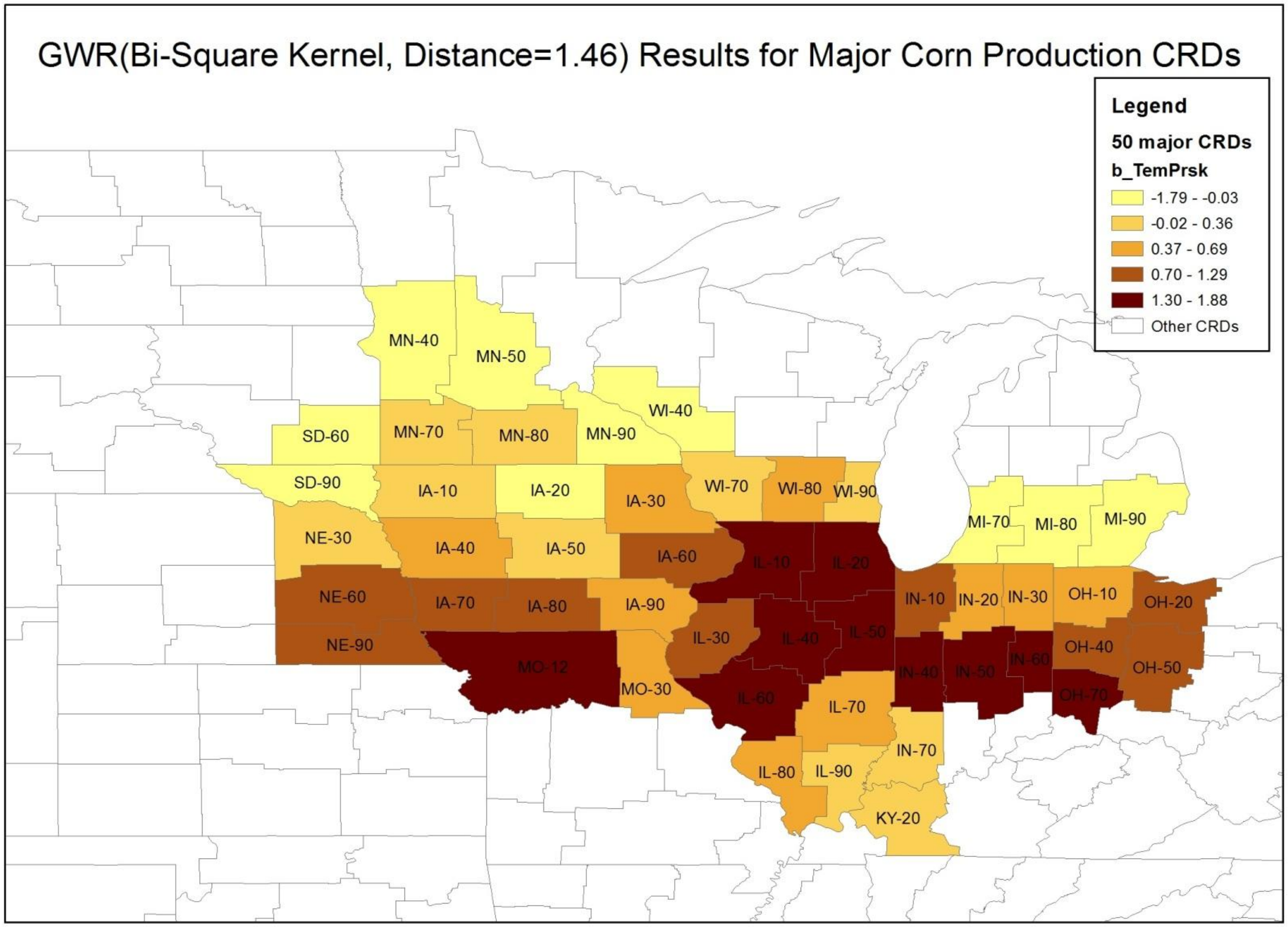


Figure 24. Coefficient Estimate of Post-silking Temperature in Panel Geographically Weighted Regression Model for Corn, 1960 – 2010

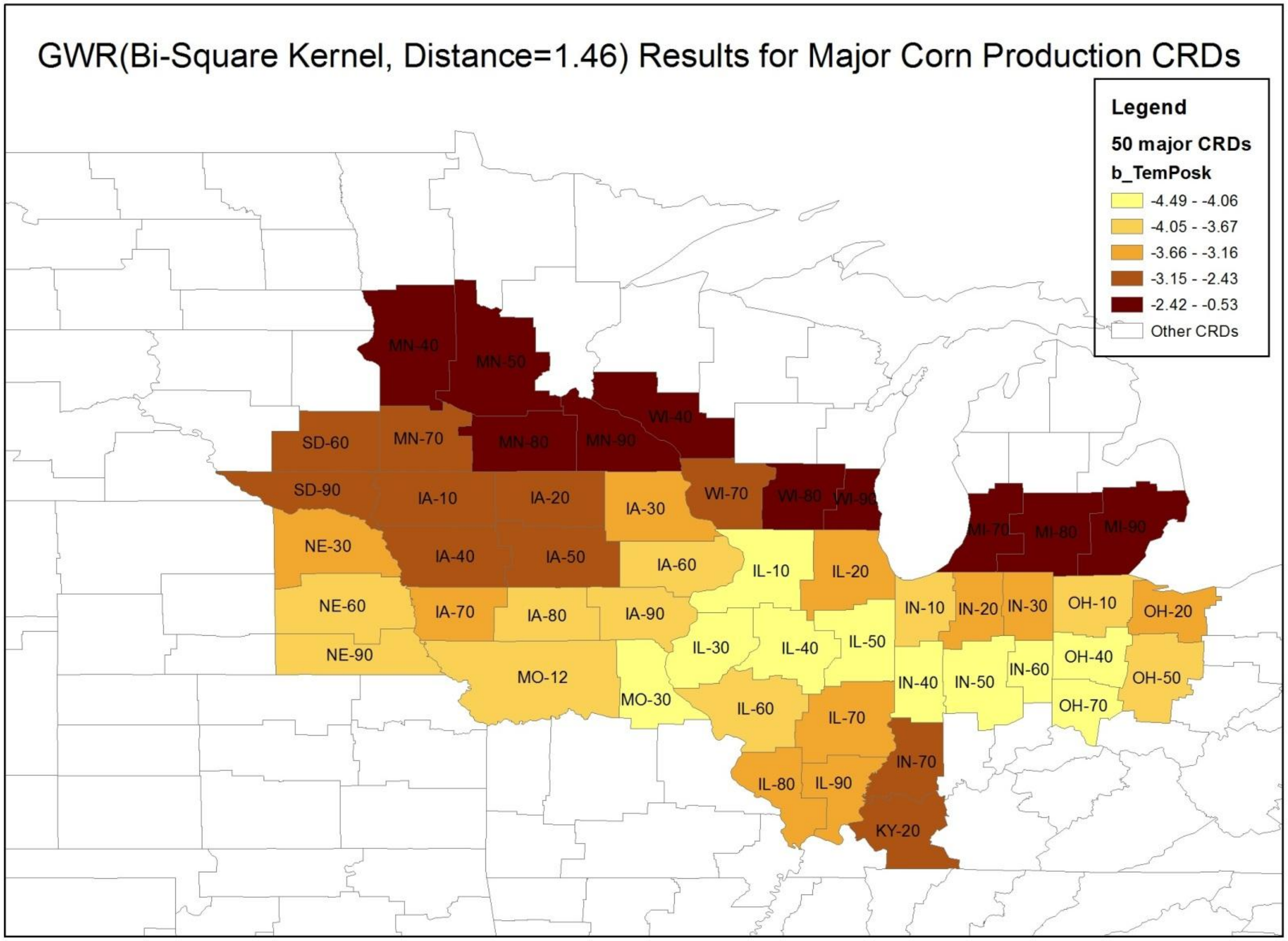


Figure 25. Coefficient Estimate of Trend in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

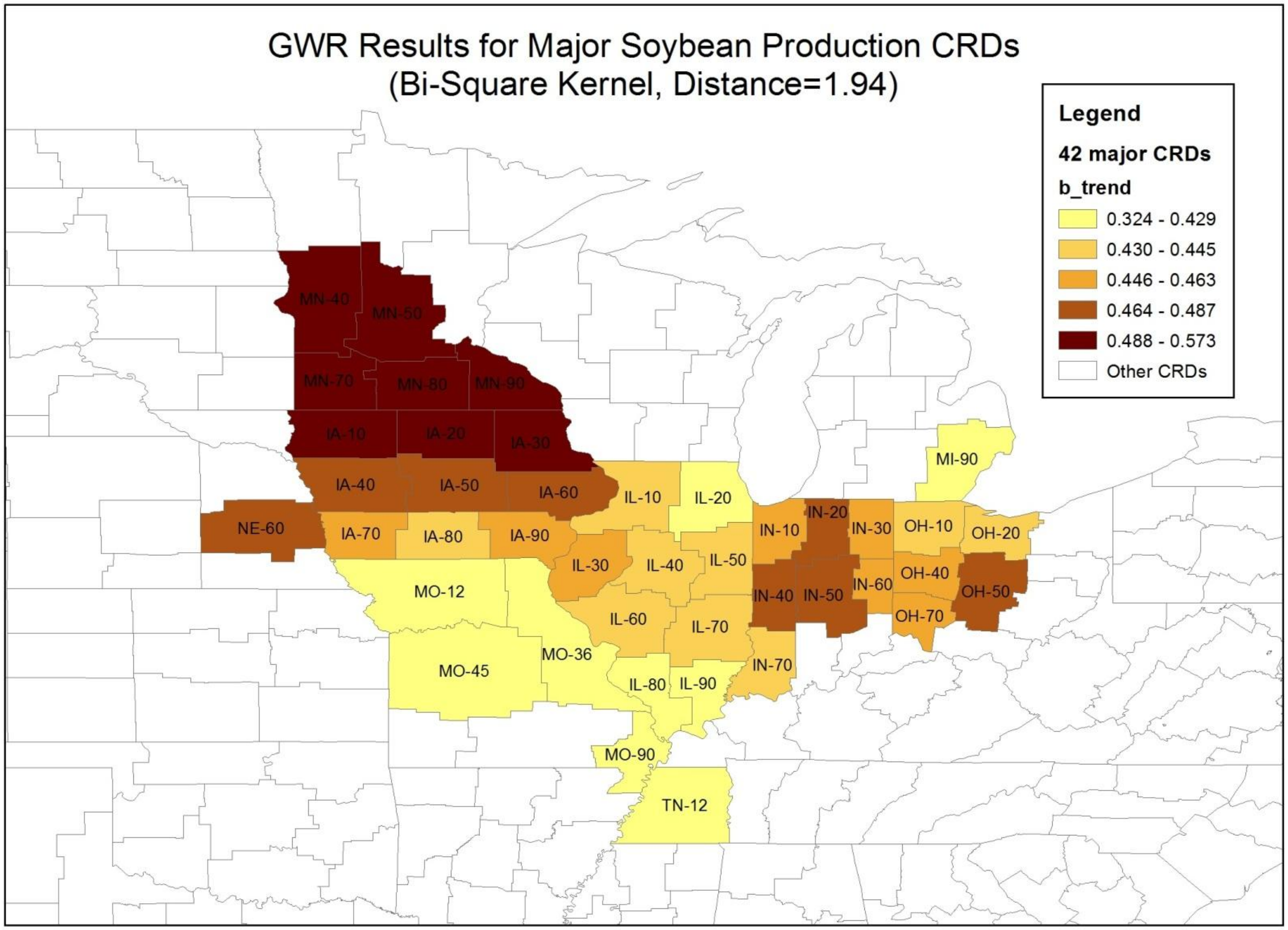


Figure 26. Coefficient Estimate of Late Planting in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

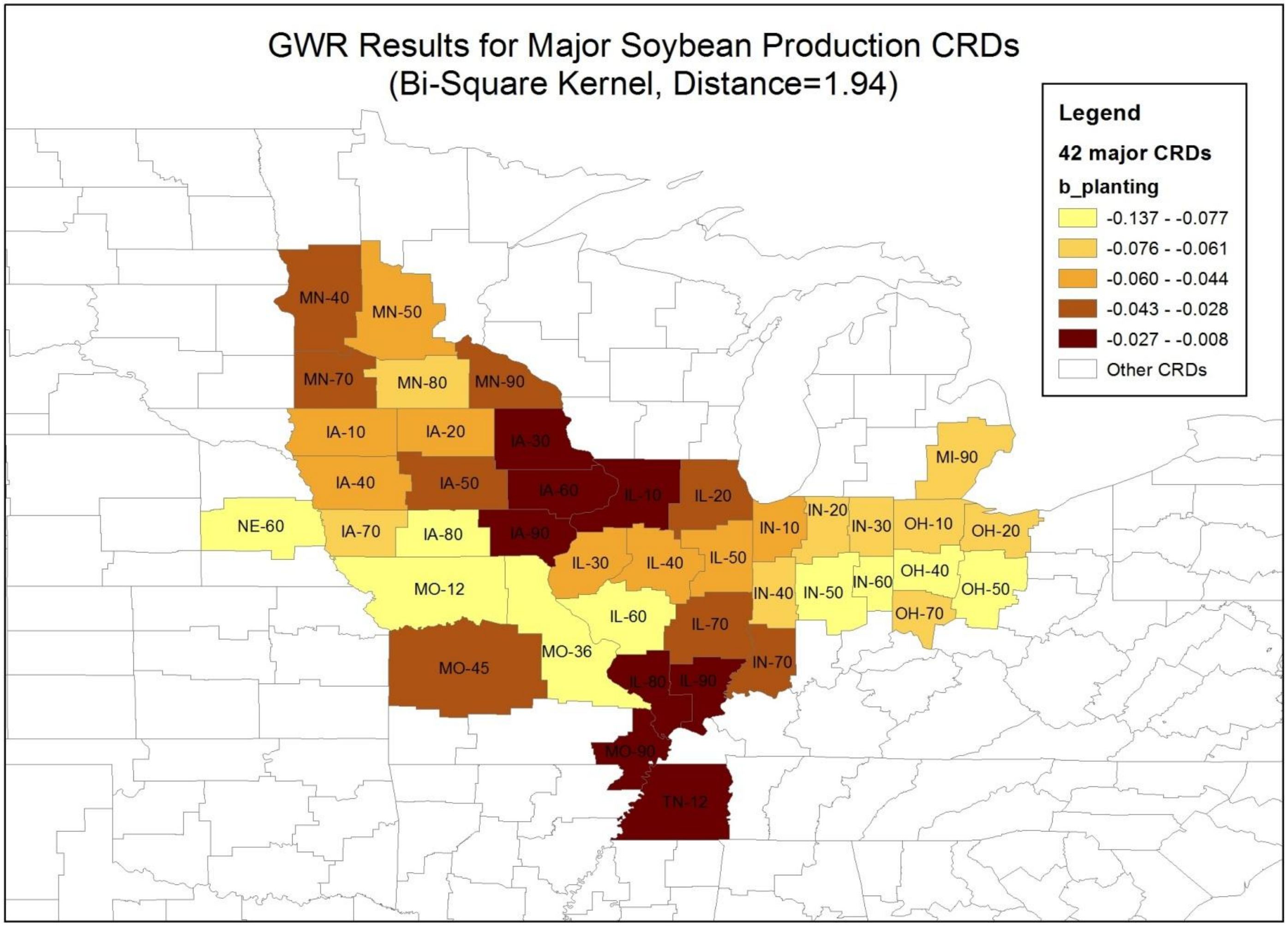


Figure 27. Marginal Effect of Pre-season Precipitation in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

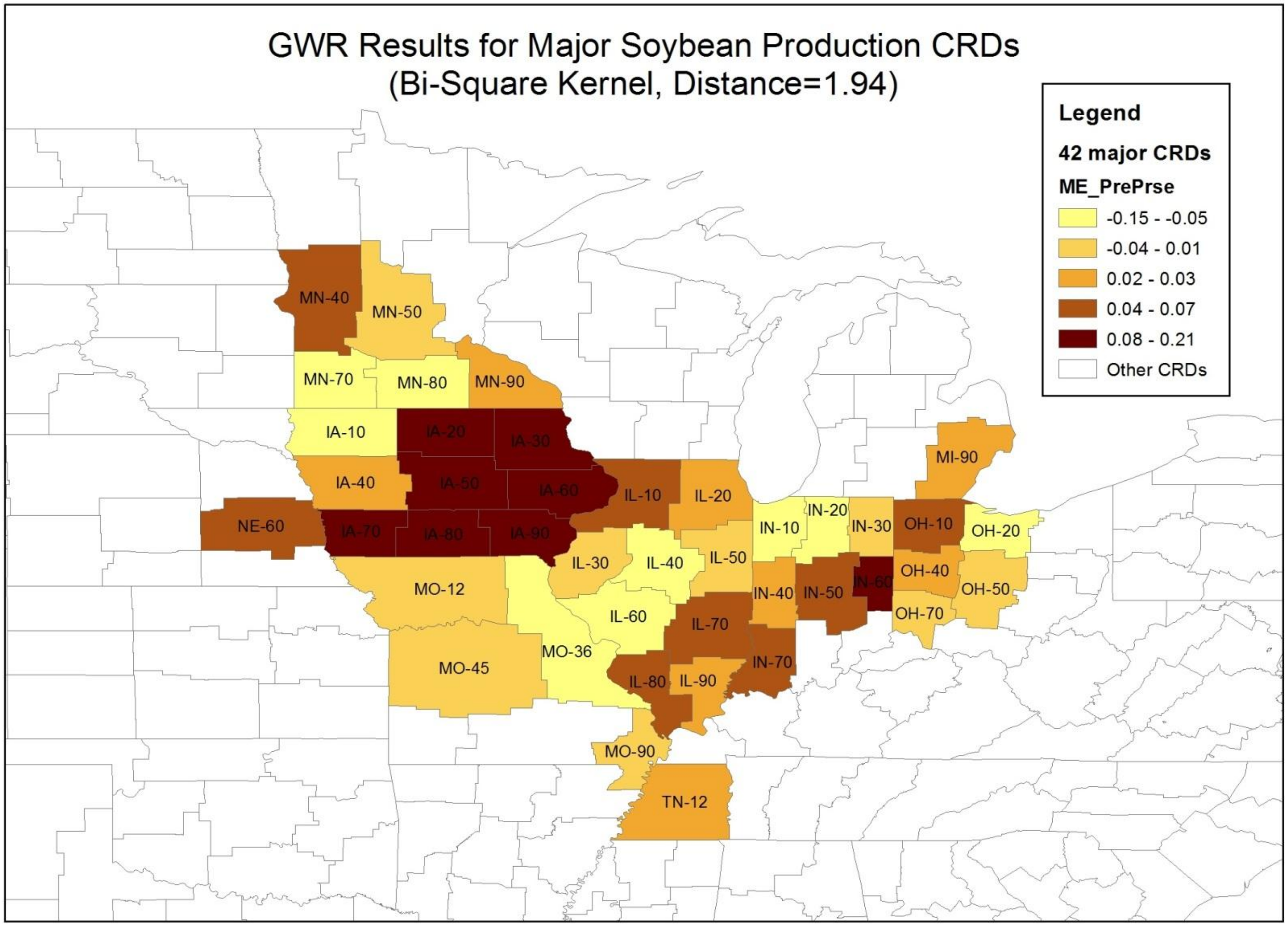


Figure 28. Marginal Effect of Pre-flowering Precipitation in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

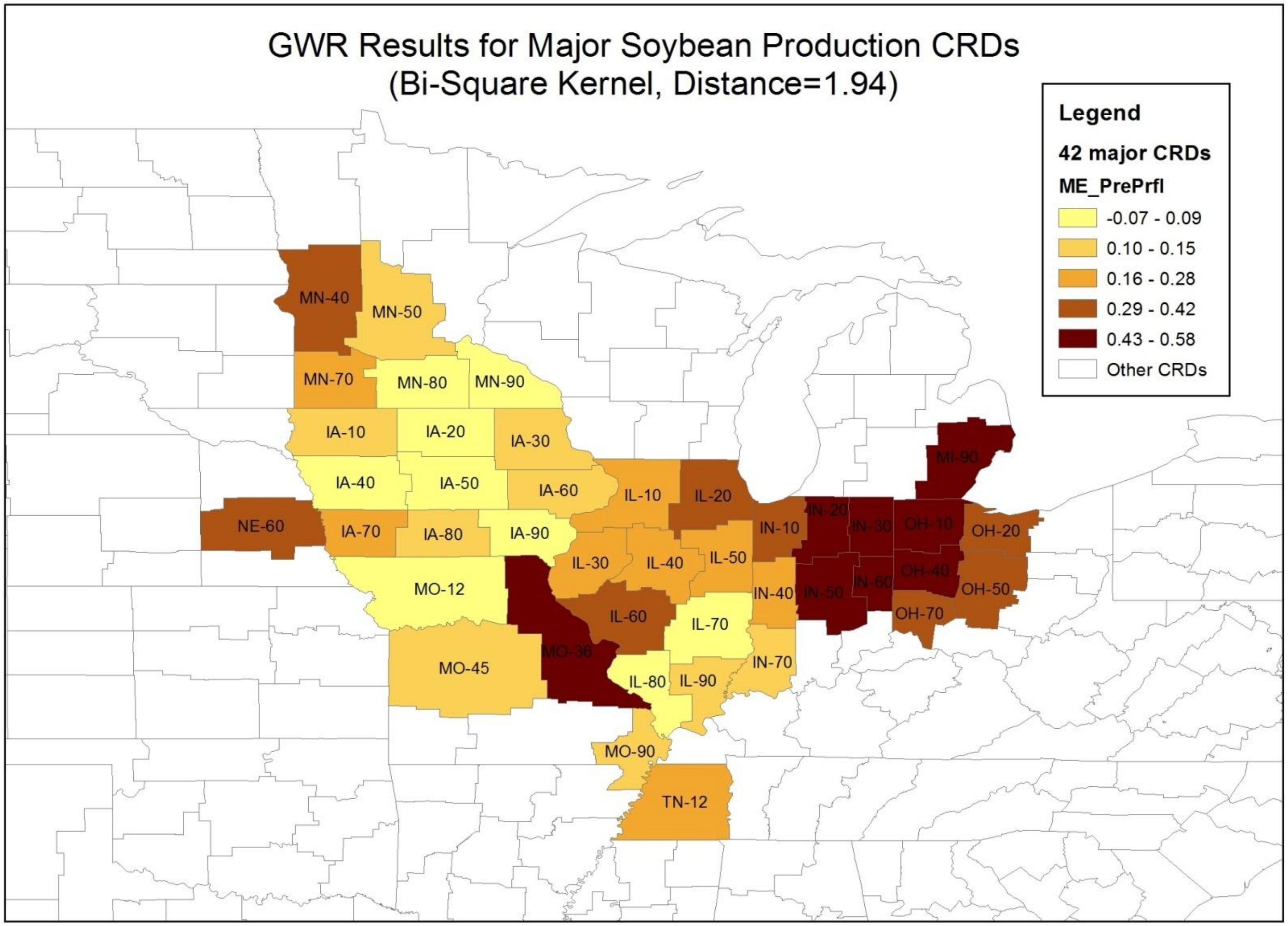


Figure 29. Marginal Effect of Post-flowering Precipitation in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

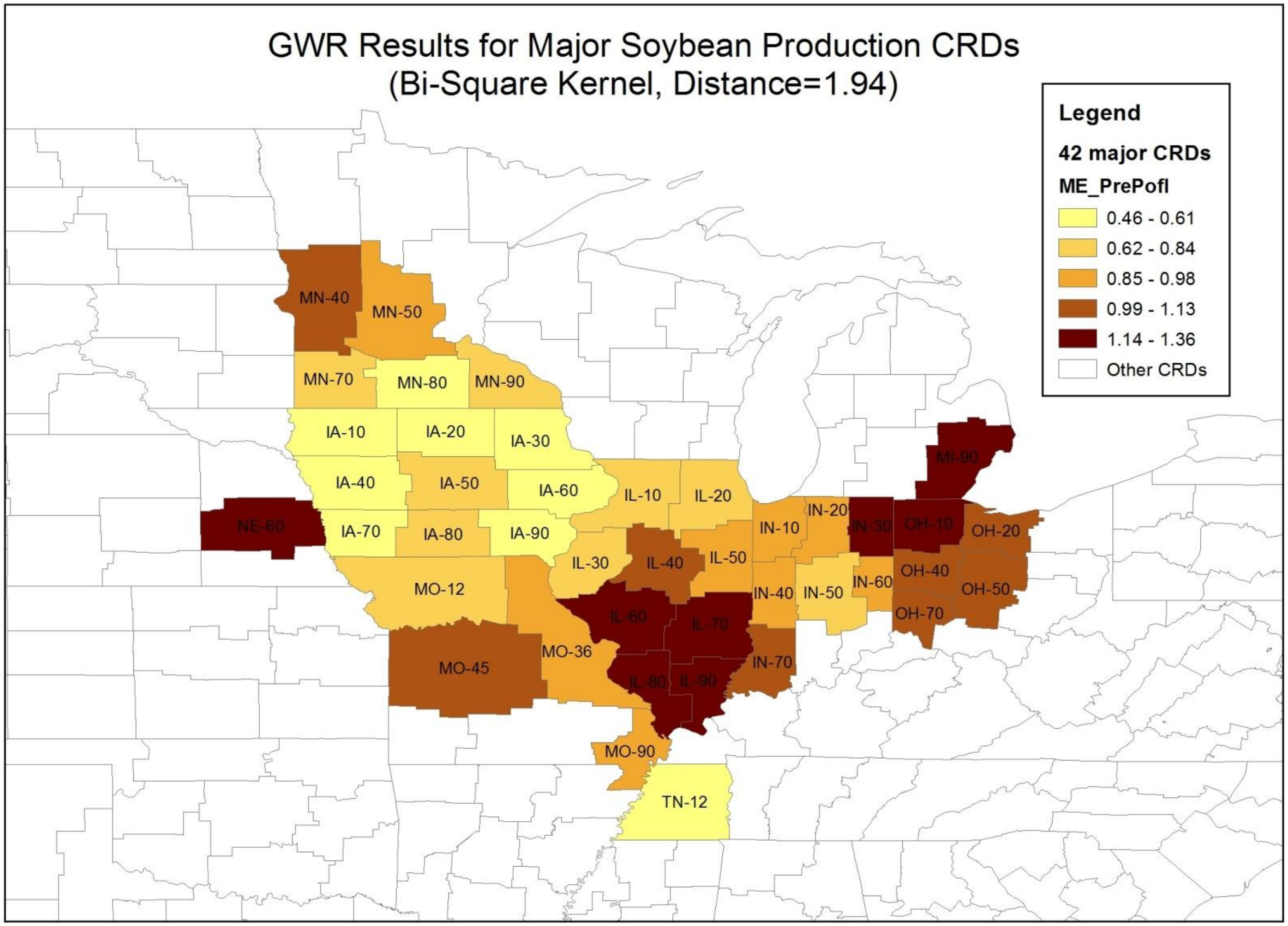


Figure 30. Coefficient Estimate of Pre-flowering Temperature in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

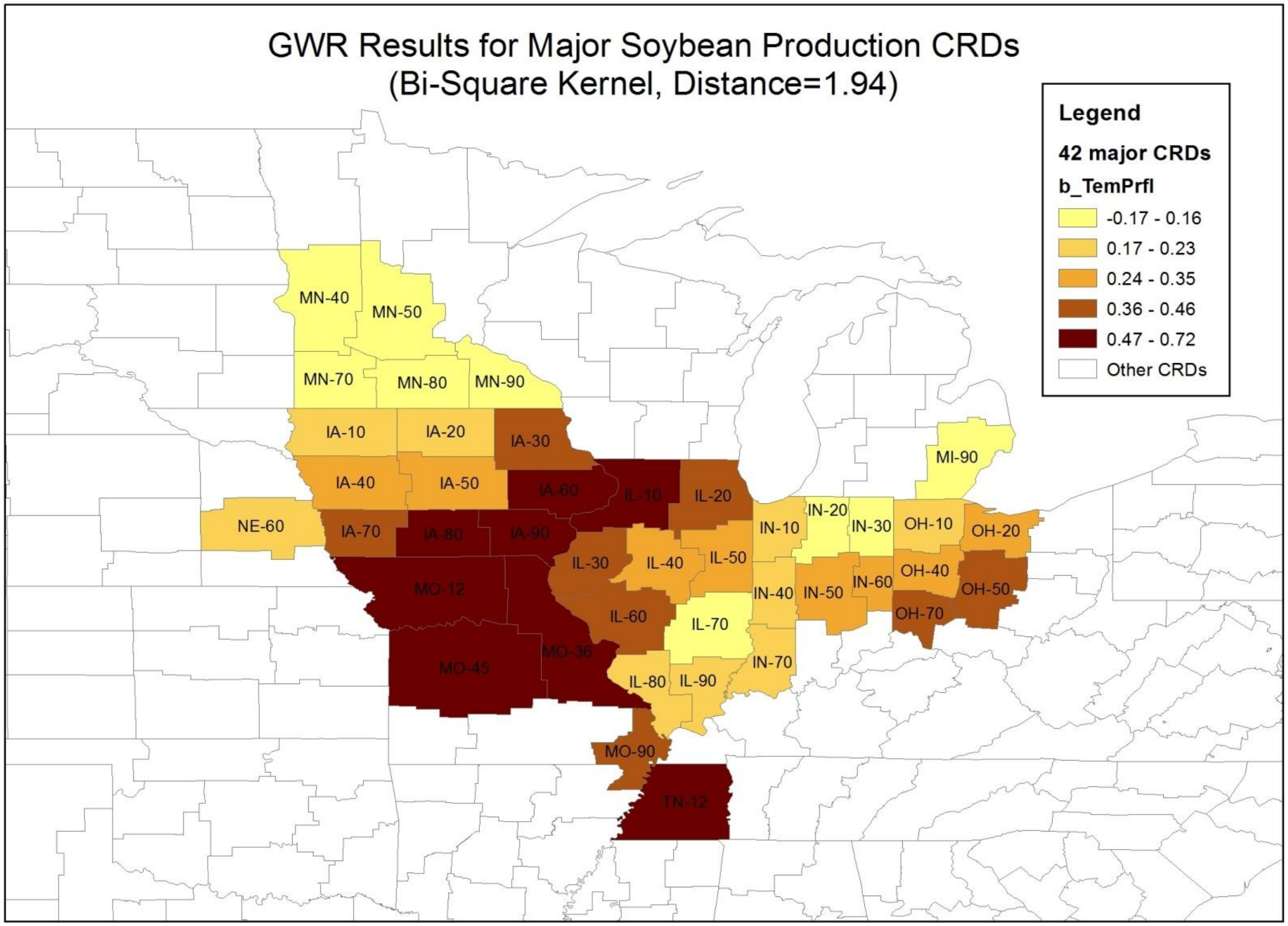


Figure 31. Coefficient Estimate of Post-flowering Temperature in Panel Geographically Weighted Regression Model for Soybeans, 1960 – 2010

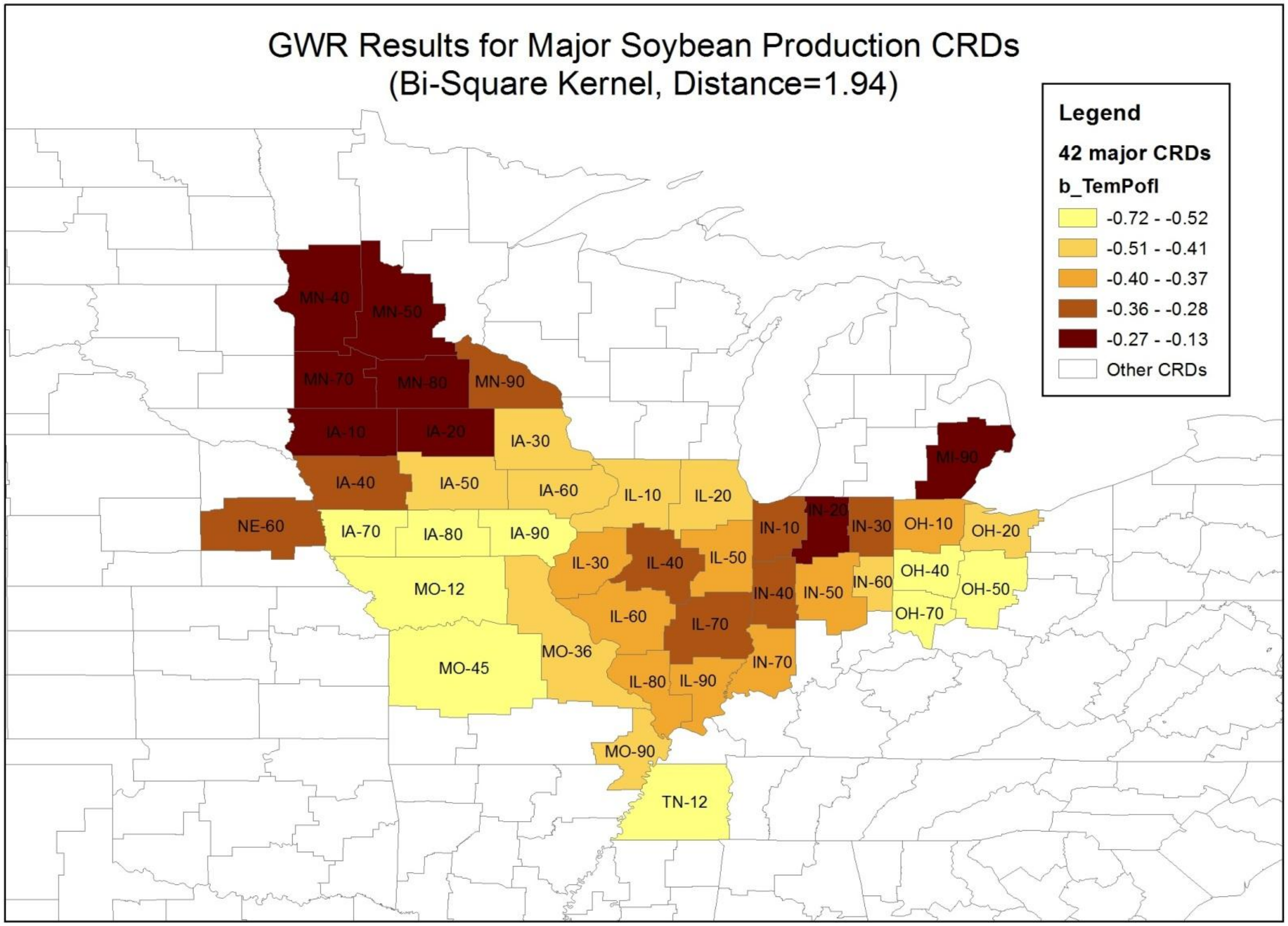


Figure 32. Out-of-Sample Forecast Errors of U.S. Total Corn Yield, 1988 – 2005/2010

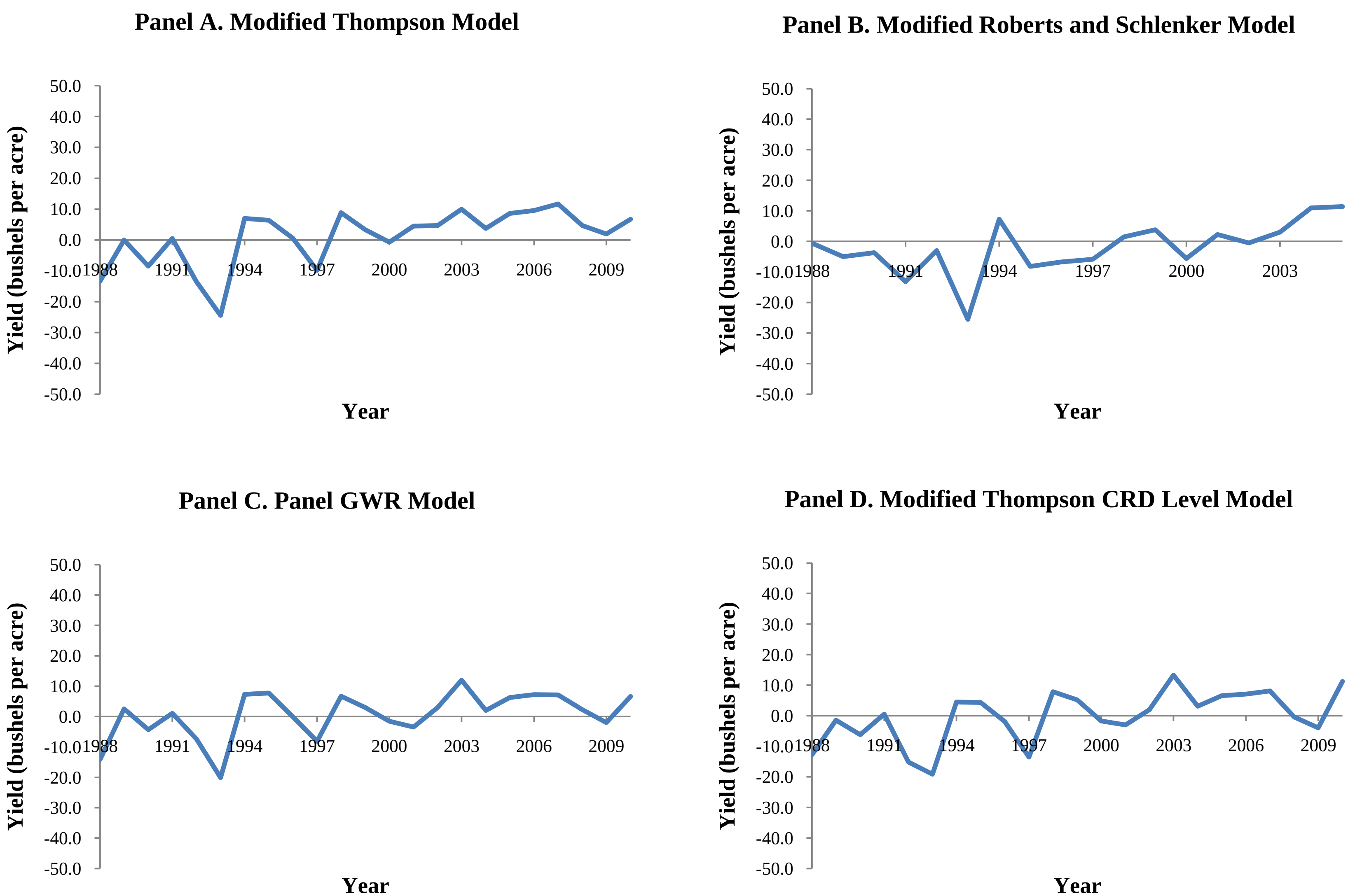
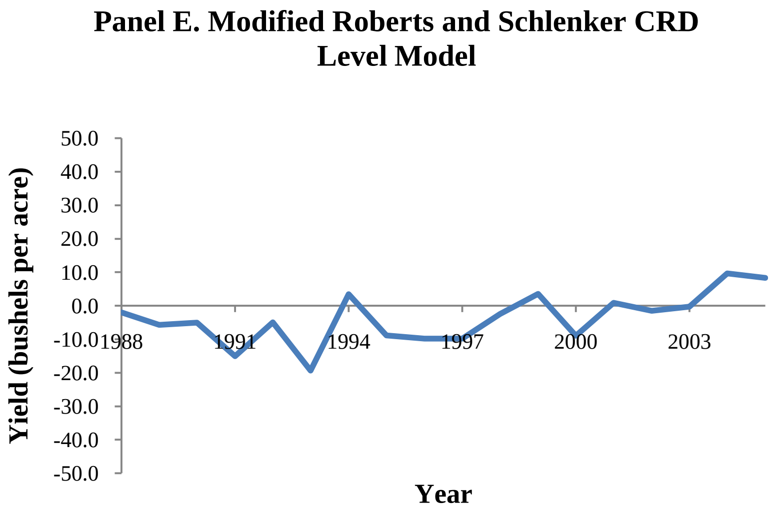
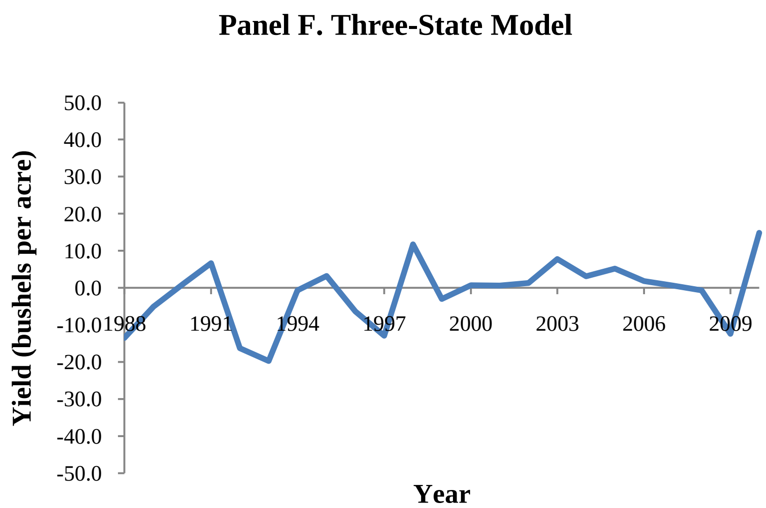
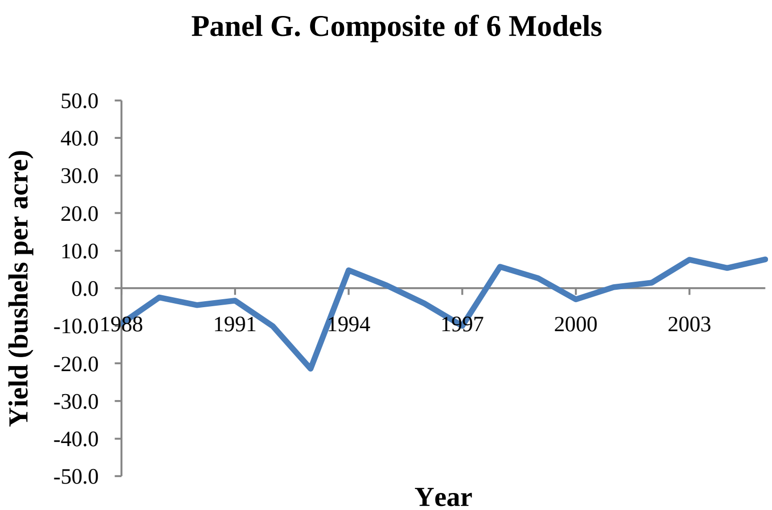


Figure 32. Out-of-Sample Forecast Errors of U.S. Total Corn Yield, 1988 – 2005/2010 (Continued)







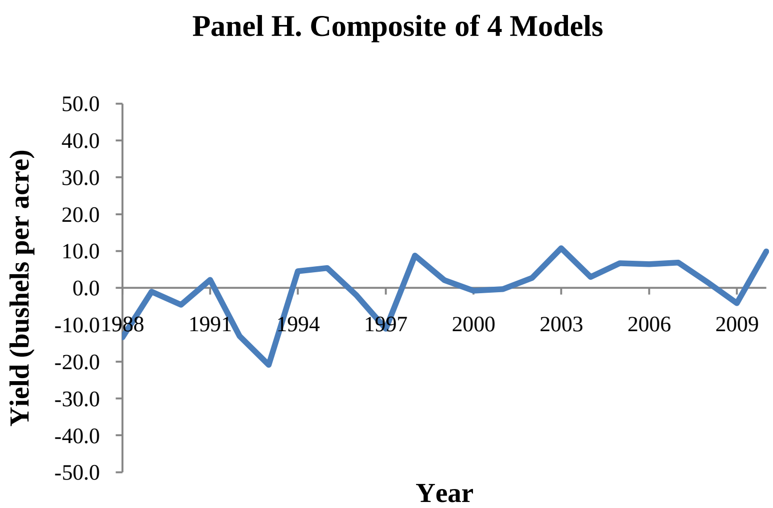


Figure 33. Out-of-Sample Forecast Errors of U.S. Total Soybean Yield, 1988 – 2005/2010

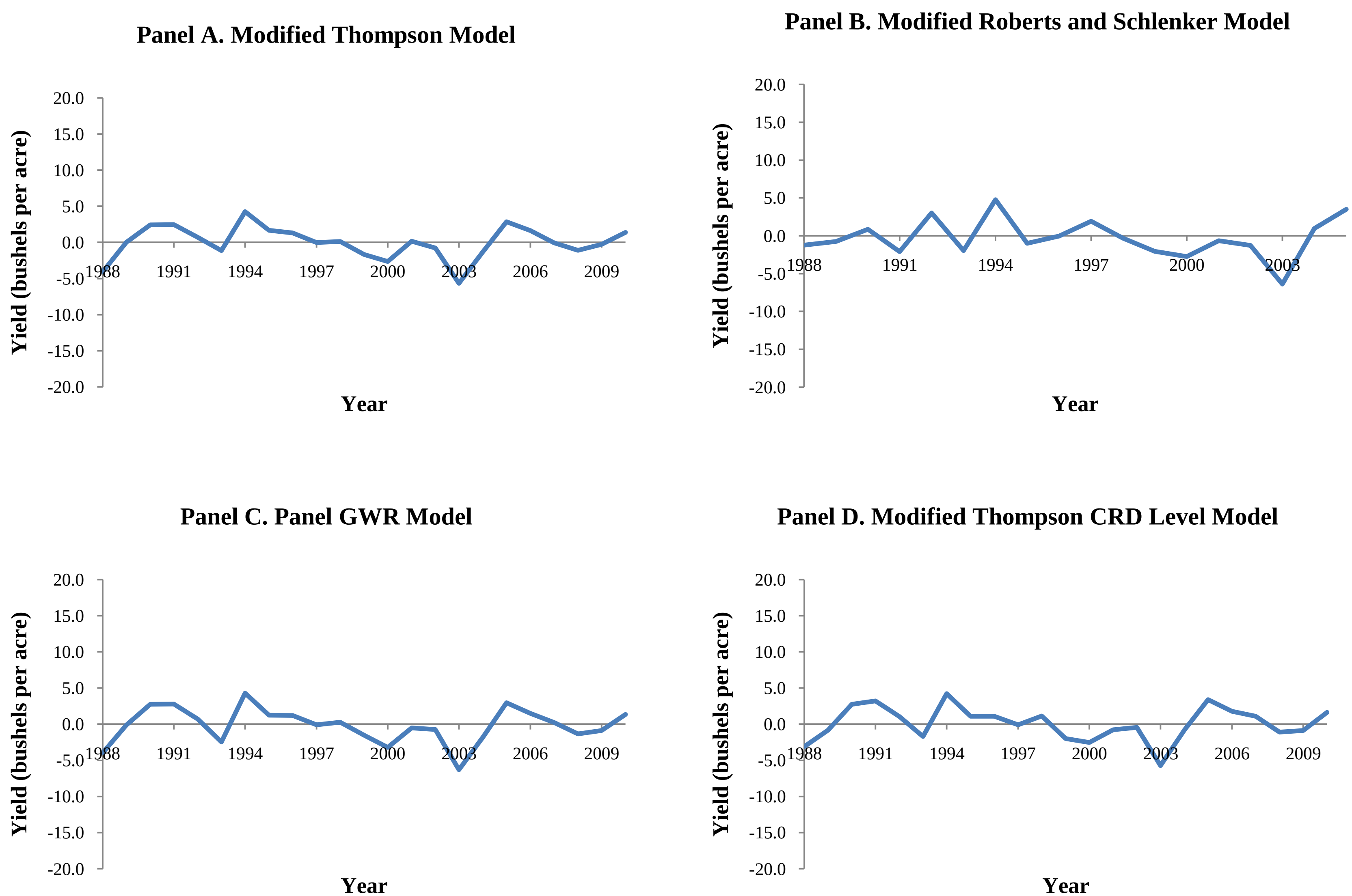
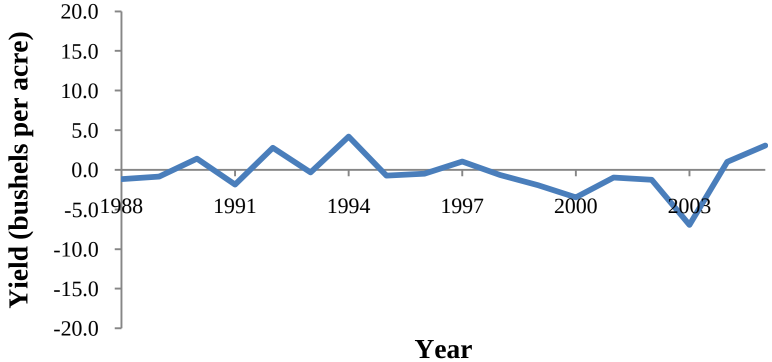
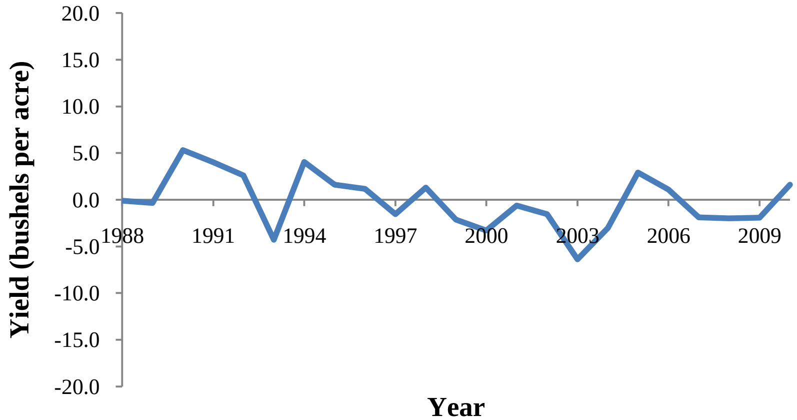


Figure 33. Out-of-Sample Forecast Errors of U.S. Total Soybean Yield, 1988 – 2005/2010 (Continued)

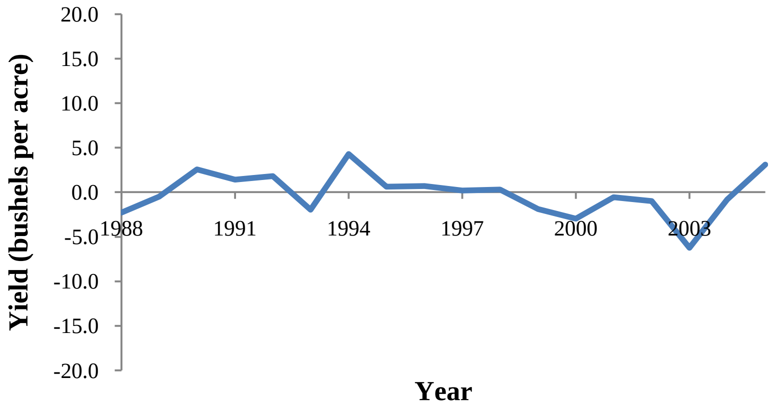
Panel E. Modified Roberts and Schlenker CRD Level Model



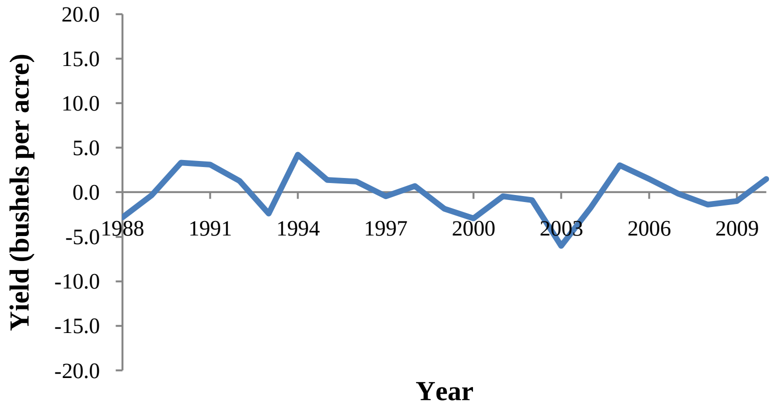
Panel F. Three-State Model



Panel G. Composite of 6 Models



Panel H. Composite of 4 Models



REFERENCES

Boote, K.J., Jones, J.W., Hoogenboom, G. "Simulation of Crop Growth: CROPGRO Model." In *Agricultural Systems Modeling and Simulation*, edited by R.M. Peart and R.B. Curry, 651-692. New York: Marcel Dekker, Inc., 1998.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussiere, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillere, J.P., Henault, C., Maraux, F., Seguin, B., Sinoquet, H. "An overview of the crop model STICS." *European Journal of Agronomy* 18, nos.3-4 (2003): 309-332.

Cassman, Kenneth G., Patricio Grassini, and Justin vanWart. "Crop Yield Potential, Yield Trends, and Global Food Security in a Changing Climate." In *Handbook of Climate Change and Agroecosystems*, edited by Daniel Hillel, and Cynthia Rosenzweig, 37-51. London: Imperial College Press, 2010.

Crop Production. U.S. Department of Agriculture, National Agricultural Statistics Service, Agricultural Statistics Board. http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1046.

Crop Progress. U.S. Department of Agriculture, National Agricultural Statistics Service, Agricultural Statistics Board. http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1048.

Fotheringham, A. Stewart, Chris Brunsdon, and Marin Charlton. *Geographically Weighted Regression*. West Sussex, England: John Wiley & Sons, Ltd., 2002.

Garcia, P., S.E. Offutt, M. Pinar, and S.A. Changnon. "Corn Yield Behavior: Effects of Technological Advance and Weather Conditions." *Journal of Climate and Applied Meteorology* 26, no.9 (1987): 1092-1102.

Good, D. L., and S. H. Irwin. "Understanding USDA Corn and Soybean Production Forecasts: Methods, Performance and Market Impacts over 1970-2005." AgMAS Project Research Report 2006-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, February 2006. http://www.farmdoc.illinois.edu/marketing/agmas/reports/06_01/AgMAS06_01.html.

Grassini, Patricio, H.S. Yang, and K.G. Cassman. "Limits to Maize Productivity in Western Corn-Belt: A Simulation Analysis for Fully Irrigated and Rainfed Conditions." *Agricultural and Forest Meteorology* 149, no.8 (2009): 1254-1265.

Irwin, Scott, D. Good, and M. Tannura. "Forming Expectations About 2008 U.S. Corn and Soybean Yields-Application of Crop Weather Models that Incorporate Planting Progress."

Marketing and Outlook Briefs 08-03, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, July 2008. http://www.farmdoc.illinois.edu/marketing/mobr/mobr_08-03/mobr_08-03.html.

Irwin, Scott, D. Good, and M. Tannura. "Early Prospects for 2009 Corn Yields in Illinois, Indiana, and Iowa." Marketing and Outlook Briefs 09-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, June 2009. http://www.farmdoc.illinois.edu/marketing/mobr/mobr_09-01/mobr_09-01.html.

Irwin, Scott, D. Good, and M. Tannura. "Early Prospects for 2009 Soybean Yields in Illinois, Indiana, and Iowa." Marketing and Outlook Briefs 09-02, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, July 2009. http://www.farmdoc.illinois.edu/marketing/mobr/mobr_09-02/mobr_09-02.html.

Kravchenko, Alexandra N., and D.G.Bullock. "Correlation of Corn and Soybean Grain Yield with Topography and Soil Properties." *Agronomy Journal* 92, no.1 (2000): 75-83.

Muchow, R.C., Sinclair, T.R., Bennett, J.M. "Temperature and Solar Radiation Effects on Potential Maize Yields across Locations." *Agronomy Journal* 82, no.2 (1990): 338-342.

Nafziger, E. "Thinking About Corn Planting Date and Population." *The Bulletin*, no.2, Article 7, University of Illinois Extension, April 4, 2008. http://www.ipm.uiuc.edu/bulletin/article.php?id=890.

National Climatic Data Center (NCDC). "Data Documentation for Data Set 9640 (DSI-9640)." NCDC Dataset Documentation & Metadata, 2002. http://www1.ncdc.noaa.gov/pub/data/documentlibrary/tddoc/td9640.pdf.

Nielsen, R.L. "Corn Planting Date is Important, But…" Corny News Network Articles, Department of Agronomy, Purdue University, April 12, 2008. http://www.kingcorn.org/news/articles.08/PltDate-0412.html.

Pecinovsky, K., and G.O. Benson. "Twenty-Six Years of Soybean Planting Date Studies." ISRFO1-13, Northeast Research and Demonstration Farms, Iowa State University, 2001. http://www.ag.iastate.edu/farms/2001reports/ne/Twentysixyearsofsoybean.pdf.

Roberts, Michael J., and Wolfram Schlenker. *Is Agricultural Production Becoming More or Less Sensitive to Extreme Heat? Evidence from US Corn and Soybean Yields.* Chicago: University of Chicago Press, 2010.

Roberts, Michael J., and Wolfram Schlenker. "The Evolution of Heat Tolerance for Corn: Implications for Climate Change." In *The Economics of Climate Change: Adaptations*

Past and Present, edited by Gary D. Libecap and Richard H. Steckel, 225-251. Chicago: University of Chicago Press, 2011.

Schlenker, Wolfram, and Michael J. Roberts. “Nonlinear Effects of Weather on Corn Yields.” *Review of Agricultural Economics* 28, no.3 (2006): 391-398.

Schlenker, Wolfram, and Michael J. Roberts. “Estimating the Impact of Climate Change on Crop Yields: The Importance of Nonlinear Temperature Effects.” NBER Working Paper No. 13799. February 2008. http://www.nber.org/papers/w13799.

Schlenker, Wolfram, and Michael J. Roberts. “Nonlinear Temperature Effects Indicate Severe Damages to US Crop Yield Under Climate Change.” *PNAS* 106, no.37 (2009): 15594-15598.

Setiyono, T.D., Cassman K.G., Specht J.E. Dobermann, A., Weiss, A., Yang, H., Conley, S.P, Robinson, A.P., Pedersen, P., De Bruin, J.L. “Simulation of Soybean Growth and Yield in Near-optimal Growth Conditions.” *Field Crops Research* 119, no.1 (2010): 161-174.

Sinclair, T.R. “Water and Nitrogen Limitation in Soybean Grain Production I. Model Development.” *Field Crops Research* 15, no.2 (1986): 125–141.

Smith, W.J. “The Effect of Weather upon the Yield of Corn.” *Monthly Weather Review* 42, no.2 (1914): 78-93.

Tannura, Michael A. “Weather, Corn and Soybean Yields, and Technology in the U.S. Corn Belt.” Master’s Thesis, University of Illinois at Urbana-Champaign, 2007.

Tannura, M.A., S. H. Irwin, and D. L. Good. “Weather, Technology, and Corn and Soybean Yields in the U.S. Corn Belt.” Marketing and Outlook Research Report 08-01, Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, February 2008. http://www.farmdoc.illinois.edu/marketing/morr/morr_archive.html.

Thompson, L.M. “An Evaluation of Weather Factors In the Production of Corn.” *The Center for Agricultural and Economic Adjustment, Iowa State University, Ames, Iowa*, Report 12T (1962).

Thompson, L.M. “Weather and Technology in the Production of Corn and Soybeans.” *The Center for Agricultural and Economic Development, Iowa State University,Ames, Iowa*, Report 17 (1963).

Thompson, L.M. “Weather and Technology in the Production of Corn in the U.S. Corn Belt.” *Agronomy Journal* 61, no.3 (1969): 453-456.

Thompson, L.M. “Weather and Technology in the Production of Soybeans in the Central U.S.” *Agronomy Journal* 62, no.2 (1970): 232-236.

Thompson, L.M. “Weather Variability, Climatic Change, and Grain Production.” *Science*188, no.4188 (1975): 535-541.

Thompson, L.M. “Weather Variability, Climatic Change, and Soybean Production.” *Journal of Soil and Water Conservation* 40, no.4 (1985): 386-389.

Thompson, L.M. “Climatic Change, Weather Variability, and Corn Production.” *Agronomy Journal* 78, no.4 (1986): 649-653.

Thompson, L.M. “Effects of Changes In Climate and Weather Variability On the Yields of Corn and Soybeans.” *Journal of Production Agriculture* 1, no.1 (1988):20-27.

Thuzar M, A. B. Puteh, N. A. P. Abdullah, M. B. Mohd. Lassim, Kamaruzaman Jusoff. “The Effects of Temperature Stress on the Quality and Yield of Soya Bean.” *Journal of Agricultural Science* 2, no.1 (2010): 172-179.

van Diepen, C.A., Wolf, J., van Keulen, H., Rappoldt, C. “WOFOST: A Simulation Model of Crop Production.” *Soil Use and Manage* 5, no.1 (1989): 16–24.

van Ittersum, M.K., Leffelaar, P.A., van Keulen, H., Kropff, M.J., Bastiaans, L., Goudriaan, J. “On Approaches An Applications of the Wageningen Crop Models.” *European Journal of Agronomy* 18, nos.3-4 (2003): 201–234.

Weekly Weather and Crop Bulletin. U.S. Department of Commerce, National Oceanic and Atmospheric Administration, and U.S. Department of Agriculture, National Agricultural Statistics Service, World Agricultural Outlook Board. http://www.usda.gov/oce/weather/pubs/Weekly/Wwcb/.

World Agricultural Supply and Demand Estimates. U.S. Department of Agriculture, National Agricultural Statistics Service, Agricultural Statistics Board. http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1194.

Yang, H.S., A. Dobermann, J.L. Lindquist, D.T. Walters, and T.J. Arkebauer. “Hybrid-Maize – A Maize Simulation Model That Combines Two Crop Modeling Approaches.” *Field Crops Research* 87, nos.2-3 (2004): 131-154.

Yu, Tian, and Bruce A. Babcock. “Are U.S. Corn and Soybeans Becoming More Drought Tolerant?” *American Journal of Agricultural Economics* 92, no.5 (2010): 1310-1323.